NONLINEAR MULTIOBJECTIVE OPTIMIZATION

INTERNATIONAL SERIES IN OPERATIONS RESEARCH & MANAGEMENT SCIENCE

Frederick S. Hillier, Series Editor
Department of Engineering-Economic Systems and
Operations Research
Stanford University
Stanford, California

Saigal, Romesh
LINEAR PROGRAMMING: A Modern Integrated Analysis

Nagurney, Anna/ Zhang, Ding
PROJECTED DYNAMICAL SYSTEMS AND VARIATIONAL INEQUALITIES WITH APPLICATIONS

Padberg, Manfred/ Rijal, Minendra P.
LOCATION, SCHEDULING, DESIGN AND INTEGER PROGRAMMING

Vanderbei, Robert J.
LINEAR PROGRAMMING: Foundations and Extensions

Jaiswal, N.K.
MILITARY OPERATIONS RESEARCH: Quantitative Decision Making

Gal, Tomas / Greenberg, Harvey J.
ADVANCES IN SENSITIVITY ANALYSIS AND PARAMETRIC PROGRAMMING

Prabhu, N.U.
FOUNDATIONS OF QUEUEING THEORY

Fang, S.-C./Rajasekera, J.R./ Tsao, H.-S.J.
ENTROPY OPTIMIZATION AND MATHEMATICAL PROGRAMMING

Yu, Gang
OPERATIONS RESEARCH IN THE AIRLINE INDUSTRY

Ho, Teck-Hua / Tang, Christopher S.
PRODUCT VARIETY MANAGEMENT

El-Taha, Muhammad / Stidham, Jr., Shaler
SAMPLE-PATH ANALYSIS OF QUEUEING SYSTEMS

NONLINEAR MULTIOBJECTIVE OPTIMIZATION

by
Kaisa Miettinen, PhD
University of Jyväskylä

Springer Science+Business Media, LLC

Library of Congress Cataloging-in-Publication Data

Miettinen, Kaisa.
Nonlinear multiobjective optimization / by Kaisa Miettinen.
p. cm. -- (International series in operations research & management science : 12)
Includes bibliographical references and index.
ISBN 978-1-4613-7544-9 ISBN 978-1-4615-5563-6 (eBook)
DOI 10.1007/978-1-4615-5563-6
1. Multiple criteria decision making. 2. Nonlinear programming.
I. Title. II. Series.
T57.95.M52 1999
658.4'03--dc21 98-37888
CIP

To my parents
Anna-Liisa and Kauko
with love and respect

CONTENTS

PREFACE

Life inevitably involves decision making, choices and searching for compromises. It is only natural to want all of these to be as good as possible, in other words, optimal. The difficulty here lies in the (at least partial) conflict between our various objectives and goals. Most everyday decisions and compromises are made on the basis of intuition, common sense, chance or all of these. However, there are areas where mathematical modelling and programming are needed, such as engineering and economics. Here, the problems to be solved vary from designing spacecraft, bridges, robots or camera lenses to blending sausages, planning and pricing production systems or managing pollution problems in environmental control. Many phenomena are of a nonlinear nature, which is why we need tools for nonlinear programming capable of handling several conflicting or incommensurable objectives. In this case, methods of traditional single objective optimization are not enough; we need new ways of thinking, new concepts, and new methods — nonlinear multiobjective optimization.

Problems with multiple objectives and criteria are generally known as *multiple criteria optimization* or *multiple criteria decision-making* (MCDM) *problems*. The area of multiple criteria decision making has developed rapidly, as the statistics collected in Steuer et al. (1996) demonstrate. For example, by the year 1994, a number of 144 conferences had been held and over 200 books and proceedings volumes had appeared on the topic. Moreover, some 1216 refereed journal articles were published between 1987 and 1992.

The MCDM field is so extensive that there is good reason to classify problems on the basis of their characteristics. They can be divided into two distinct types (in accordance with MacCrimmon (1973)). Depending on the properties of the feasible solutions, we distinguish multiattribute decision analysis and multiobjective optimization. In *multiattribute decision analysis*, the set of feasible alternatives is discrete, predetermined and finite. Specific examples are the selection of the locations of power plants and dumping sites or the purchase of cars and houses. In *multiobjective optimization* problems, the feasible alternatives are not explicitly known in advance. An infinite number of them exists and they are represented by decision variables restricted by constraint functions. These problems can be called continuous. In these cases, one has to generate the alternatives before they can be valuated.

As far as multiattribute decision analysis is concerned, we refer to the monographs by Hwang and Yoon (1981) and Keeney and Raiffa (1976). More ref-

erences, together with 17 major methods in the area accompanied by simple examples, can be found in the latter monograph. A more recent summary of the methodology is given in Yoon and Hwang (1995). A brief historical account, including the basic ideas behind both multiobjective optimization and multiattribute decision analysis together with suggestions for further reading, can be found in Dyer et al. (1992) and Zionts (1992). (The latter also handles multiattribute utility theory and negotiation.) In addition, a review of the research in both of these problem classes accompanied by future directions appears in Korhonen et al. (1992a). It contains short descriptions of many concepts and areas in multiple criteria optimization and decision making not included here.

In this book we concentrate solely on continuous multiobjective optimization. This does not mean that some of the methods presented cannot be applied to multiattribute decision analysis. Nevertheless, most of the methods have been designed only for one or other of the problem types, exploiting certain special characteristics.

The importance of multiobjective optimization can be seen from the large variety of applications presented in the literature. Some idea of its possibilities can be gained from the fact that over 500 papers describing different applications (between the years 1955 and 1986) are listed in White (1990). They cover, for example, problems concerning agriculture, banking, the health service, energy, industry, water and wildlife.

Even though we have restricted ourselves to handling only multiobjective optimization problems, it nonetheless remains a broad area of research and we are therefore obliged to omit several topics to be able to give a uniform presentation. We shall restrict the treatment to deterministic problems. Nevertheless, a few words and further references are in order in relation to problems involving uncertainties. These can be divided into stochastic and fuzzy problems. In stochastic programming it is usually assumed that uncertainty is due to a lack of information about prevailing states, and that this uncertainty only concerns the occurrence of the states and not the definition of the states, results or criteria themselves. A problem containing random variables as coefficients on a certain probability space is called a stochastic programming problem (treated, for example, in the monographs of Guddat et al. (1985) and Stancu-Minasian (1984)). When decision making takes place in an environment where the goals, constraints and consequences of possible actions are not precisely known, it is called decision making in fuzzy environments (handled, for example, in the proceedings of Kacprzyk and Orlovski (1987)). Fuzzy coefficients may also be involved in the problem formulation. Both stochastic and fuzzy multiobjective optimization (for linear problems) are dealt with and compared in the proceedings of Slowinski and Teghem (1990). Let us stress once again that here we assume the problems to be deterministic; that is, the outcome of any feasible solution is known for certain.

Solving problems with several conflicting objectives usually requires the participation of a human decision maker who can express preference relations

between alternative solutions and who continues from the point where mathematical tools end. Here we assume that a single decision maker is involved. With several decision makers, the whole question of problem setting is very different. In addition to the mathematical side of the solution process, there is also the aspect of negotiation and consensus striving between the decision makers. The number of decision makers affects the means of approaching and solving the problem significantly. A summary of group decision making is given in the monograph of Hwang and Lin (1987). Here we settle for one decision maker.

A number of specific problem types requires special handling (not included here). Among these are problems in which the feasible solutions must have integer values or 0-1 values, multiobjective trajectory optimization problems, where the multiple objectives have multiple observation points, multiobjective networks or transportation networks and multiobjective dynamic programming. Here we shall not go into these areas but adhere to standard methods.

Thus far we have outlined our interest here as being in deterministic continuous multiobjective optimization with a single decision maker. This definition still contains two broad areas, namely linear and nonlinear cases. Because linear programming utilizes the special characteristics of the problem, its methods are not usually applicable to nonlinear problems. Further, linear multiobjective optimization theory and methodology have been extensively treated in the literature, so there is no reason to repeat them here. One of the best presentations focusing mainly on linear problems is Steuer (1986). However, the methodology of nonlinear multiobjective optimization has not been drawn together since Hwang and Masud (1979) (currently out of print). One more fact to notice is that improved computational capacity enables problems to be handled without linearizations and simplifications. Finally, linear problems are a subset of nonlinear problems and that is why nonlinear methods can be used in both cases. For these reasons, this book concentrates on nonlinear multiobjective optimization.

The aim here is to provide an up-to-date, self-contained and consistent survey and review of the literature and the state of the art on nonlinear (deterministic) multiobjective optimization starting with basic results.

The amount of literature on multiobjective optimization is immense. The treatment in this book is based on about 1500 publications in English printed mainly after the year 1980. Almost 700 of them are cited and listed in the bibliography. This extensive list of references supplements the contents regarding areas not covered.

Problems related to real-life applications often contain irregularities and nonsmoothnesses. The treatment of nondifferentiable multiobjective optimization in the literature is rather rare. For this reason we also include in this book material about the possibilities, background, theory and methods of nondifferentiable multiobjective optimization.

Theory and methods for multiobjective optimization have been developed chiefly during the last four decades. Here we do not go into the history as

the origin and the achievements in this field of research from 1776 to 1960 are widely treated in Stadler (1979). A brief summary of the history is also given in Gal and Hanne (1997). There it is demonstrated that multiobjective optimization has its foundations in utility theory and economics, game theory, mathematical research on order relations and vector norms, linear production theory, and nonlinear programming.

Let us mention some further readings. The monographs of Chankong and Haimes (1983b), Cohon (1978), Hwang and Masud (1979), Osyczka (1984), Sawaragi et al. (1985), Steuer (1986) and Yu (1985) provide an extensive overview of the area of multiobjective optimization. Further noteworthy monographs on the topic are those of Rietveld (1980), Vincke (1992) and Zeleny (1974, 1982). A significant part of Vincke (1992) deals, however, with multiattribute decision analysis. The behavioural aspects of multiobjective optimization are mostly treated in Ringuest (1992), whereas the theoretical aspects are extensively handled in the monographs by Jahn (1986a) and Luc (1989).

As far as this book is concerned, the contents are divided into three parts. Part I provides the theoretical background. Chapter 1 leads into the topic and Chapter 2 presents important notation, concepts and definitions in multiobjective optimization with some illustrative figures. Various theoretical aspects appear in Chapter 3. For example, analogous optimality conditions for differentiable and nondifferentiable problems are considered. A solid, conceptual basis and foundation for the remainder of the book is laid. Throughout the book we keep to problems involving only finite-dimensional Euclidean spaces. (Dauer and Stadler (1986) provide a survey on multiobjective optimization in infinite-dimensional spaces.)

The methodology is handled in Part II. Methods are divided into four classes in Chapter 1 according to the role of a (single) decision maker in the solution process. The state of the art in method development is portrayed by describing a number of different methods accompanied by their theoretical background in Chapters 2 to 5. For ease of comparison, all the methods are presented using a uniform notation. The good and the weak properties of the methods are also introduced with references to extensions and applications. The class of interactive methods in Chapter 5 contains most of the methods, and it is the most extensively handled. Linear problems and methods are only occasionally touched on. In addition to describing solution methods, we introduce some implementations. In connection with every method described, some author's comments appear in the concluding remarks. Some of the methods are depicted in more detail and some only mentioned. Appropriate references to the literature are always included.

Part III is Related Issues. After the presentation of a set of different solution methods, some comparison is appropriate in Chapter 1. Naturally, no absolute order of superiority can be given, but some points can be raised. A table comparing some of the features of the interactive methods described is included. In addition, we present brief summaries of some of the comparisons

available in the literature. Moreover, we suggest some outlines regarding the important question of selecting an appropriate method. Method selection itself is a problem with multiple objectives. Nevertheless, in addition to considering some significant factors, we present a decision tree to aid selection. This tree contains all the interactive methods previously described in some detail. It is based on some of the fundamental assumptions underlying the methods and different ways of exchanging information between the method and its user.

Compared with the plethora of methods, only a relatively few computer implementations are widely known and available. However, some implementational aspects are touched on and some software mentioned in Chapter 2.

As computers and monitors have developed, graphical illustration has increased in importance and has also become easier to produce. Hence graphical illustration of alternative solutions together with related matters are featured in Chapter 3. The potential and restrictions of graphics are treated and some clarifying figures are enclosed.

We conclude with comments on future directions in Chapter 4 and an epilogue in Chapter 5.

This book is intended both for researchers and students in the areas such as (applied) mathematics, engineering, economics, operations research and management science; it is meant both for professionals and practitioners in many different fields of application. For beginners, this book provides an introduction to the theory and methodology of nonlinear multiobjective optimization. For other readers, it offers an extensive reference to many related results and methods. Obviously it is not possible in a single book to include all the aspects and methods of nonlinear multiobjective optimization. However, the intention has been to provide a consistent summary using a uniform notation leading to further references. The uniform style of presentation may help in selecting an appropriate method for the problem to be solved. It is hoped the extensive bibliography will be of value to researchers.

The book gives sufficient theoretical background to allow those interested to follow the derivation of the featured methods. However, the theoretical treatment in Chapter 3 of Part I, for example, is not essential for the continuation. For both theoretically and practically oriented readers, the algorithms are described in a consistent manner with some implementational remarks and software information also presented. Because, however, this is not an actual textbook, no exercises or illustrative examples have been included.

NOTATION AND SYMBOLS

$\mathbf{R}^n$	n-dimensional Euclidean space	
f_i	objective function	page 5
k	number of objective functions	page 5
$\mathbf{x}$	decision (variable) vector	page 5
n	number of decision variables	page 5
S	feasible region	page 5
Z	feasible objective region	page 5
$\mathbf{f}(\mathbf{x})$, $\mathbf{z}$	objective vector	page 5
$\mathbf{R}^n_+$	nonnegative orthant of $\mathbf{R}^n$	page 6
$\|\mathbf{x}\|$	Euclidean norm	page 6
$\text{dist}(\mathbf{x}, S)$	Euclidean distance function	page 6
$B(\mathbf{x}, \delta)$	open ball	page 6
$\text{int}\, S$	interior of set S	page 6
$\text{conv}\, S$	convex hull of set S	page 6
$\nabla f_i(\mathbf{x})$	gradient of f_i at $\mathbf{x}$	Definition 2.1.4, page 7
$\frac{\partial f_i(\mathbf{x})}{\partial x_j}$	partial derivative of f_i subject to x_j	Definition 2.1.4, page 7
$\mathbf{R}^n_\varepsilon$	blunt cone in $\mathbf{R}^n$	page 9
$\nabla^2 f_i(\mathbf{x})$	Hessian matrix of f_i at $\mathbf{x}$	Definition 2.1.11, page 9
$\partial f_i(\mathbf{x})$	subdifferential of f_i at $\mathbf{x}$	Definition 2.1.14, page 10
$\boldsymbol{\xi}$	subgradient	Definition 2.1.14, page 10
$\bar{z}_i$	aspiration level	Definition 2.3.1, page 14
$\bar{\mathbf{z}}$	reference point	Definition 2.3.1, page 14
$\mathbf{z}^\star$	ideal objective vector	Definition 2.4.1, page 16
$\mathbf{z}^{\star\star}$	utopian objective vector	Definition 2.4.2, page 16
$\mathbf{z}^{\text{nad}}$	nadir objective vector	page 16
U	value function	Definition 2.6.1, page 21
D	ordering cone	page 23
λ_{ij}	trade-off rate	Definition 2.8.4, page 27
Λ^G_{ij}	global trade-off	Definition 2.8.5, page 27
m_{ij}	marginal rate of substitution	Definition 2.8.6, page 28
$\boldsymbol{\lambda}$, $\boldsymbol{\mu}$	Karush-Kuhn-Tucker multipliers	page 38
$s_{\bar{\mathbf{z}}}$	achievement (scalarizing) function	page 108
P	number of alternative objective vectors	

ACKNOWLEDGEMENTS

I wish to express my gratitude to several individuals. First, I must thank Professor Pekka Neittaanmäki, who originally proposed multiobjective optimization to me as a research subject. I also want to take this opportunity to express a special thank you to Professor Pekka Korhonen whose ideas and contacts helped me in publishing this book.

I am indebted to lecturers Michael Freeman, Ari Lehtonen and Jukka-Pekka Santanen as well as to my colleague Marko Mäkelä for their efforts in reading the manuscript and suggesting improvements. In addition, I wish to thank Marko for availing me freely of his expertise in nondifferentiable optimization.

I am grateful to my colleague Timo Männikkö for his helpful TEXnical hints. Further, I wish to express my appreciation to Markku Könkkölä as well as Tapani Tarvainen for their help in numerous practical and technical matters in the course of the preparation of this book, not to mention Markku's efforts with some of the figures. I also want to thank my students for their assistance in the implementations and Marja-Leena Rantalainen for some revisions.

On this occasion a special vote of thanks should be extended to all those software developers who have given their programs to me for test purposes.

My thanks go to the Academy of Finland (grant number 22346) for financial support and to the University of Jyväskylä for providing me with the facilities in which to work.

Finally, I want to truly acknowledge the support and love of my dear parents, Anna-Liisa and Kauko, who laid the foundations for my education.

Kaisa Miettinen

Part I

TERMINOLOGY AND THEORY

1. INTRODUCTION

We begin by laying a conceptual and theoretical basis for the continuation and restrict our treatment to finite-dimensional Euclidean spaces. First, we present the deterministic, continuous problem formulation to be handled and some general notation. Then we introduce several concepts and definitions of multiobjective optimization as well as their interconnections. The concepts and terms used in the field of multiobjective optimization are not completely fixed. The terminology used here is occasionally slightly different from that in general use. In some cases, only one of the existing terms is employed. Somewhat different definitions of concepts are presented, for example, in Zionts (1989).

To deepen the theoretical basis, we treat optimality conditions for differentiable and nondifferentiable multiobjective optimization problems. We also briefly touch on the topics of sensitivity analysis, stability and duality.

Throughout the book, even some simple results are proved, for the convenience of the reader (with possible appropriate references), in order to lay firm cornerstones for the continuation. However, to keep the text to a reasonable length, some proofs have been omitted if they can directly be found as such elsewhere. In those cases, appropriate references in the literature are indicated.

Multiobjective optimization problems are usually solved by scalarization. *Scalarization* means that the problem is converted into a single (scalar) or a family of single objective optimization problems. In this way the new problem has a real-valued objective function, possibly depending on some parameters. After the multiobjective optimization problem has been scalarized, the widely developed theory and methods for single objective optimization can be used. Even though multiobjective optimization methods are presented in Part II, we emphasize here at the outset that the methods and the theory of single objective optimization are presumed to be known.

2. CONCEPTS

This chapter introduces the basic concepts of (nonlinear) multiobjective optimization and the notations used in the continuation.

2.1. Problem Setting and General Notation

We begin by defining the problem to be handled.

2.1.1. Multiobjective Optimization Problem

We study a *multiobjective optimization problem* of the form

$$\begin{aligned} &\text{minimize} && \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\} \\ &\text{subject to} && \mathbf{x} \in S, \end{aligned} \tag{2.1.1}$$

where we have k (≥ 2) *objective functions* $f_i\colon \mathbf{R}^n \to \mathbf{R}$. We denote the vector of objective functions by $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}))^T$. The *decision (variable) vectors* $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ belong to the (nonempty) *feasible region* (set) S, which is a subset of the *decision variable space* $\mathbf{R}^n$. We do not yet fix the form of the *constraint functions* forming S, but refer to S in general.

The word 'minimize' means that we want to minimize all the objective functions simultaneously. If there is no conflict between the objective functions, then a solution can be found where every objective function attains its optimum. In this case, no special methods are needed. To avoid such trivial cases we assume that there does not exist a single solution that is optimal with respect to every objective function. This means that the objective functions are at least partly conflicting. They may also be incommensurable (i.e., in different units).

In the following, we denote the image of the feasible region by Z ($= \mathbf{f}(S)$) and call it a *feasible objective region*. It is a subset of the *objective space* $\mathbf{R}^k$. The elements of Z are called *objective (function) vectors* or *criterion vectors* and denoted by $\mathbf{f}(\mathbf{x})$ or $\mathbf{z} = (z_1, z_2, \ldots, z_k)^T$, where $z_i = f_i(\mathbf{x})$ for all $i = 1, \ldots, k$ are *objective (function) values* or *criterion values*. The words in the parentheses above are usually omitted for short.

K. Miettinen, *Nonlinear Multiobjective Optimization*

For clarity and simplicity of the treatment we assume that all the objective functions are to be minimized. If an objective function f_i is to be maximized, it is equivalent to minimize the function $-f_i$.

In what follows, whenever we refer to a multiobjective optimization problem, it is problem (2.1.1) unless stated otherwise. Finding a solution to (2.1.1) in one way or another is called a *solution process* in the continuation.

2.1.2. Background Concepts

First, we present some general concepts and notations. We use bold face and superscripts for vectors, for example, $\mathbf{x}^1$, and subscripts for components of vectors, for example, x_1. All the vectors here are assumed to be column vectors. For two vectors, $\mathbf{x}$ and $\mathbf{x}^* \in \mathbf{R}^n$, the notation $\mathbf{x}^T\mathbf{x}^*$ denotes their *scalar product* and the vector inequality $\mathbf{x} \leq \mathbf{x}^*$ means that $x_i \leq x_i^*$ for all $i = 1, \ldots, n$. Correspondingly $\mathbf{x} < \mathbf{x}^*$ stands for $x_i < x_i^*$ for all $i = 1, \ldots, n$.

The nonnegative orthant of $\mathbf{R}^n$ is denoted by $\mathbf{R}^n_+$. In other words, $\mathbf{R}^n_+ = \{\mathbf{x} \in \mathbf{R}^n \mid x_i \geq 0 \text{ for } i = 1, \ldots, n\}$. The *Euclidean norm* of a vector $\mathbf{x} \in \mathbf{R}^n$ is denoted by $\|\mathbf{x}\| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$. The *Euclidean distance function* between a point $\mathbf{x}^*$ and a set S is denoted by $\operatorname{dist}(\mathbf{x}^*, S) = \inf_{\mathbf{x} \in S} \|\mathbf{x}^* - \mathbf{x}\|$. The symbol $B(\mathbf{x}^*, \delta)$ denotes an *open ball* with a centre $\mathbf{x}^*$ and a radius $\delta > 0$, $B(\mathbf{x}^*, \delta) = \{\mathbf{x} \in \mathbf{R}^n \mid \|\mathbf{x}^* - \mathbf{x}\| < \delta\}$. The notation $\operatorname{int} S$ stands for the *interior* of a set S.

The vectors $\mathbf{x}^i$, $i = 1, \ldots, m$, are *linearly independent* if the only weighting coefficients β_i for which $\sum_{i=1}^m \beta_i \mathbf{x}^i = 0$ are $\beta_i = 0$, $i = 1, \ldots, m$. The sum $\sum_{i=1}^m \beta_i \mathbf{x}^i$ is called a *convex combination* of the vectors $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^m \in S$, if $\beta_i \geq 0$ for all i and $\sum_{i=1}^m \beta_i = 1$. The *convex hull* of a set $S \subset \mathbf{R}^n$, denoted by $\operatorname{conv} S$, is the set of all convex combinations of vectors in S.

A set $S \subset \mathbf{R}^n$ is a *cone* if $\beta\mathbf{x} = (\beta x_1, \ldots, \beta x_m)^T \in S$ whenever $\mathbf{x} \in S$ and $\beta \geq 0$. The negative of a cone is $-S = \{-\mathbf{x} \in \mathbf{R}^n \mid \mathbf{x} \in S\}$. A cone S is said to be *pointed* if it satisfies $S \cap -S = \{\mathbf{0}\}$. A cone $-S$ transformed to $\mathbf{x}^* \in \mathbf{R}^n$ is denoted by $\mathbf{x}^* - S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{x} = \mathbf{x}^* + \mathbf{d}, \text{ where } \mathbf{d} \in -S\}$.

It is said that $\mathbf{d} \in \mathbf{R}^n$ is a *feasible direction* emanating from $\mathbf{x} \in S$ if there exists $\alpha^* > 0$ such that $\mathbf{x} + \alpha\mathbf{d} \in S$ for $0 \leq \alpha \leq \alpha^*$.

In some connections we assume that the feasible region is formed of inequality constraints, that is, $S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}$. An inequality constraint g_j is said to be *active* at a point $\mathbf{x}^*$ if $g_j(\mathbf{x}^*) = 0$, and the set of active constraints at $\mathbf{x}^*$ is denoted by $J(\mathbf{x}^*) = \{j \in \{1, \ldots, m\} \mid g_j(\mathbf{x}^*) = 0\}$.

Different types of multiobjective optimization problems can be defined.

Definition 2.1.1. When all the objective functions and the constraint functions forming the feasible region are linear, then the multiobjective optimization problem is called *linear*. In brief, it is an *MOLP* (multiobjective linear programming) problem.

If at least one of the objective or the constraint functions is nonlinear, the problem is called a *nonlinear multiobjective optimization problem*.

A large variety of solution techniques have been created as to enable the special characteristics of MOLP problems to be taken into account. Here we concentrate on cases where nonlinear functions are included and thus methods for nonlinear problems are needed. Methods and details of MOLP problems are mentioned only incidentally.

Before we define convex multiobjective optimization problems, we briefly write down the definitions of convex functions and convex sets.

Definition 2.1.2. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *convex* if for all $\mathbf{x}^1, \mathbf{x}^2 \in \mathbf{R}^n$ is valid that $f_i(\beta\mathbf{x}^1 + (1-\beta)\mathbf{x}^2) \leq \beta f_i(\mathbf{x}^1) + (1-\beta) f_i(\mathbf{x}^2)$ for all $0 \leq \beta \leq 1$.

A set $S \subset \mathbf{R}^n$ is *convex* if $\mathbf{x}^1, \mathbf{x}^2 \in S$ implies that $\beta\mathbf{x}^1 + (1-\beta)\mathbf{x}^2 \in S$ for all $0 \leq \beta \leq 1$.

Definition 2.1.3. The multiobjective optimization problem is *convex* if all the objective functions and the feasible region are convex.

A convex multiobjective optimization problem is an important concept in the continuation. We shall also need related generalized concepts, quasiconvex and pseudoconvex functions. The pseudoconvexity of a function calls for differentiability. For completeness, we write down the definitions of differentiable and continuously differentiable functions.

Definition 2.1.4. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *differentiable* at $\mathbf{x}^*$ if

$$f_i(\mathbf{x}^* + \mathbf{d}) - f_i(\mathbf{x}^*) = \nabla f_i(\mathbf{x}^*)^T \mathbf{d} + \|\mathbf{d}\|\, \varepsilon(\mathbf{x}^*, \mathbf{d}),$$

where $\nabla f_i(\mathbf{x}^*)$ is the *gradient* of f_i at $\mathbf{x}^*$ and $\varepsilon(\mathbf{x}^*, \mathbf{d}) \to 0$ as $\|\mathbf{d}\| \to 0$.

In addition, f_i is *continuously differentiable* at $\mathbf{x}^*$ if all of its *partial derivatives* $\frac{\partial f_i(\mathbf{x}^*)}{\partial x_j}$ $(j = 1, \ldots, n)$, that is, all the components of the gradient are continuous at $\mathbf{x}^*$.

The gradient of f_i at $\mathbf{x}^*$ can also be denoted by $\nabla_x f_i(\mathbf{x}^*)$ to emphasize that the derivation is carried out subject to $\mathbf{x}$.

Now we can define quasiconvex and pseudoconvex functions.

Definition 2.1.5. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *quasiconvex* if $f_i(\beta\mathbf{x}^1 + (1-\beta)\mathbf{x}^2) \leq \max[f_i(\mathbf{x}^1), f_i(\mathbf{x}^2)]$ for all $0 \leq \beta \leq 1$ and for all $\mathbf{x}^1, \mathbf{x}^2 \in \mathbf{R}^n$.

Let f_i be differentiable at every $\mathbf{x} \in \mathbf{R}^n$. Then it is *pseudoconvex* if for all $\mathbf{x}^1, \mathbf{x}^2 \in \mathbf{R}^n$ such that $\nabla f_i(\mathbf{x}^1)^T(\mathbf{x}^2 - \mathbf{x}^1) \geq 0$, we have $f_i(\mathbf{x}^2) \geq f_i(\mathbf{x}^1)$.

As far as the relations of quasiconvex and pseudoconvex functions are concerned, every pseudoconvex function is also quasiconvex.

The definition of convex functions can be modified for *concave* functions by replacing '$\leq$' by '$\geq$'. Correspondingly, the definition of quasiconvex functions becomes appropriate for *quasiconcave* functions by the exchange of '$\leq$' to '$\geq$' and 'max' to 'min'. In the definition of pseudoconvex functions we replace

'$\geq$' by '$\leq$' to get the definition for *pseudoconcave* functions. Notice that if a function f_i is quasiconvex, all of its *level sets* $\{\mathbf{x} \in \mathbf{R}^n \mid f_i(\mathbf{x}) \leq \alpha\}$ are convex and if it is quasiconcave, all of its level sets $\{\mathbf{x} \in \mathbf{R}^n \mid f_i(\mathbf{x}) \geq \alpha\}$ are convex (see, for example, Mangasarian (1969, pp. 133–134)).

Sometimes we also need strict definitions.

Definition 2.1.6. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *strictly convex* if $f_i(\beta\mathbf{x}^1 + (1-\beta)\mathbf{x}^2) < \beta f_i(\mathbf{x}^1) + (1-\beta)f_i(\mathbf{x}^2)$ and *strictly quasiconvex* if $f_i(\beta\mathbf{x}^1 + (1-\beta)\mathbf{x}^2) < \max[f_i(\mathbf{x}^1), f_i(\mathbf{x}^2)]$ for all $0 < \beta < 1$ and for all $\mathbf{x}^1, \mathbf{x}^2 \in \mathbf{R}^n$, where $f_i(\mathbf{x}^1) \neq f_i(\mathbf{x}^2)$.

Notice that strict convexity of a function implies convexity and convexity implies both strict quasiconvexity and quasiconvexity. If differentiability is assumed, convexity implies pseudoconvexity which implies strict quasiconvexity. See Bazaraa et al. (1993, pp. 78–118) or Mangasarian (1969, pp. 131–147) for the details of the relations. The corresponding results are valid for concave functions and their generalizations. It is worth pointing out that convexity, concavity and related concepts can be defined in a convex set $S \subset \mathbf{R}^n$ as well as in $\mathbf{R}^n$.

We also need other function types. The first of these are related to monotonicity.

Definition 2.1.7. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *increasing* if for $\mathbf{x}^1$ and $\mathbf{x}^2 \in \mathbf{R}^n$

$$x_j^1 \leq x_j^2 \quad \text{for all} \quad j = 1, \ldots, n \quad \text{imply} \quad f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2).$$

Correspondingly, the function f_i is *decreasing* if $f_i(\mathbf{x}^1) \geq f_i(\mathbf{x}^2)$.

A function is *monotonic* (or order preserving) if it is either increasing or decreasing. Monotonicity can be tightened up in several ways.

Definition 2.1.8. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *strictly increasing* if for $\mathbf{x}^1$ and $\mathbf{x}^2 \in \mathbf{R}^n$

$$x_j^1 < x_j^2 \quad \text{for all} \quad j = 1, \ldots, n \quad \text{imply} \quad f_i(\mathbf{x}^1) < f_i(\mathbf{x}^2).$$

Definition 2.1.9. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *strongly increasing* if for $\mathbf{x}^1$ and $\mathbf{x}^2 \in \mathbf{R}^n$

$x_j^1 \leq x_j^2$ for all $j = 1, \ldots, n$ and $x_l^1 < x_l^2$ for some l imply $f_i(\mathbf{x}^1) < f_i(\mathbf{x}^2)$.

Correspondingly, the function f_i is *strongly decreasing* if $f_i(\mathbf{x}^1) > f_i(\mathbf{x}^2)$.

Notice that if a function is strongly decreasing and differentiable, all of its partial derivatives have to be (strictly) negative.

In the next definition we need a subset $\mathbf{R}_e^n$ of $\mathbf{R}^n$. It is defined as

$$\mathbf{R}^n_\varepsilon = \{\mathbf{x} \in \mathbf{R}^n \mid \operatorname{dist}(\mathbf{x}, \mathbf{R}^n_+) \leq \varepsilon \|\mathbf{x}\|\}.$$

Definition 2.1.10. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *ε-strongly increasing* if for $\mathbf{x}^1$ and $\mathbf{x}^2 \in \mathbf{R}^n$

$$\mathbf{x}^1 \in \mathbf{x}^2 - \mathbf{R}^n_\varepsilon \setminus \{\mathbf{0}\} \text{ imply } f_i(\mathbf{x}^1) < f_i(\mathbf{x}^2).$$

For the convenience of the reader we define twice differentiable functions and some related concepts.

Definition 2.1.11. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *twice-differentiable* at $\mathbf{x}^*$ if

$$f_i(\mathbf{x}^* + \mathbf{d}) - f_i(\mathbf{x}^*) = \nabla f_i(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \nabla^2 f_i(\mathbf{x}^*)\mathbf{d} + \|\mathbf{d}\|^2 \varepsilon(\mathbf{x}^*, \mathbf{d}),$$

where $\nabla f_i(\mathbf{x}^*)$ is the gradient, the symmetric $n \times n$ matrix $\nabla^2 f_i(\mathbf{x}^*)$ is a *Hessian matrix* of f_i at $\mathbf{x}^*$ and $\varepsilon(\mathbf{x}^*, \mathbf{d}) \to 0$ as $\|\mathbf{d}\| \to 0$. The Hessian matrix of a twice-differentiable function consists of *second-order partial derivatives* $\frac{\partial^2 f_i(\mathbf{x}^*)}{\partial x_j \partial x_l}$ $(j,\ l = 1, \ldots, n)$. In other words,

$$\nabla^2 f_i(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial^2 f_i(\mathbf{x}^*)}{\partial x_1^1} & \cdots & \frac{\partial^2 f_i(\mathbf{x}^*)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f_i(\mathbf{x}^*)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f_i(\mathbf{x}^*)}{\partial x_n^2} \end{pmatrix}.$$

In addition, f_i is *twice continuously differentiable* at $\mathbf{x}^*$ if all of its second-order partial derivatives are continuous at $\mathbf{x}^*$.

A symmetric $n \times n$ matrix M is called *positive definite*, if $\mathbf{x}^T M \mathbf{x} > 0$ for all $\mathbf{0} \neq \mathbf{x} \in \mathbf{R}^n$.

We shall also handle nondifferentiable multiobjective optimization problems. For that reason we define locally Lipschitzian functions (see Clarke (1983, pp. 9–11) and Mäkelä and Neittaanmäki (1992, pp. 5–10)).

Definition 2.1.12. A function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is *locally Lipschitzian* at a point $\mathbf{x}^* \in \mathbf{R}^n$ if there exist scalars $K > 0$ and $\delta > 0$ such that

$$|f_i(\mathbf{x}^1) - f_i(\mathbf{x}^2)| \leq K\|\mathbf{x}^1 - \mathbf{x}^2\| \text{ for all } \mathbf{x}^1, \mathbf{x}^2 \in B(\mathbf{x}^*, \delta).$$

Notice that a convex function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ is for any point $\mathbf{x} \in \mathbf{R}^n$ locally Lipschitzian at $\mathbf{x}$.

In what follows, a function is called *nondifferentiable* if it is locally Lipschitzian (and not necessarily continuously differentiable).

Definition 2.1.13. The multiobjective optimization problem is *nondifferentiable* if some of the objective functions or the constraint functions forming the feasible region are nondifferentiable.

According to Rademacher's Theorem (see, e.g., Federer (1969)), we know that a locally Lipschitzian function, defined in an open set, is differentiable almost everywhere in that set. A set where a function f_i is not differentiable is denoted here by Ω_{f_i}. In the sequel, we employ the concept subdifferential as defined in Clarke (1983). It corresponds to the gradient in the differentiable case.

Definition 2.1.14. Let the function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ be locally Lipschitzian at a point $\mathbf{x}^* \in \mathbf{R}^n$. The set

$$\partial f_i(\mathbf{x}^*) = \operatorname{conv}\{\boldsymbol{\xi} \in \mathbf{R}^n \mid \boldsymbol{\xi} = \lim_{l\to\infty} \nabla f_i(\mathbf{x}^l);\ \mathbf{x}^l \to \mathbf{x}^*,\ \mathbf{x}^l \in \mathbf{R}^n \setminus \Omega_{f_i}\}$$

is called a *subdifferential* of the function f_i evaluated at the point $\mathbf{x}^*$. In addition, the vectors $\boldsymbol{\xi} \in \partial f_i(\mathbf{x}^*)$ are called *subgradients*.

We end with a special type of upper semidifferentiable function (see Wang (1989)).

Definition 2.1.15. Let the function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ be locally Lipschitzian at a point $\mathbf{x}^* \in \mathbf{R}^n$. Then it is *upper semidifferentiable* at $\mathbf{x}^*$ if for every $\mathbf{d} \in \mathbf{R}^n$, any sequence $\{t_j\}_{j=1}^{\infty}$ with $t_j \to 0$ and sequence $\{\boldsymbol{\xi}^j\}$, where $\boldsymbol{\xi}^j \in \partial f_i(\mathbf{x}^* + t_j\mathbf{d})$ for every j, we have

$$\liminf_{j\to\infty} \frac{f_i(\mathbf{x}^* + t_j\mathbf{d}) - f_i(\mathbf{x}^*)}{t_j} \leq \limsup_{j\to\infty} (\boldsymbol{\xi}^j)^T\mathbf{d}.$$

Special properties of nondifferentiable functions are introduced in Section 3.2, in the context where nondifferentiability is handled.

After these general definitions and concepts we can continue with multiobjective optimization terminology.

2.2. Pareto Optimality

In this section, we handle a crucial concept in optimization, namely optimality. In single objective optimization problems, the main focus is on the decision variable space. In the multiobjective context we are often more interested in the objective space. For one thing, it is usually of a lower dimension than the decision variable space. Further, objective values are used below in defining optimality.

Because of the contradiction and possible incommensurability of the objective functions, it is not possible to find a single solution that would be optimal for all the objectives simultaneously. Multiobjective optimization problems are in a sense ill-defined. There is no natural ordering in the objective space because it is only partially ordered (meaning that, for example, $(1,1)^T$ can be said to be less than $(3,3)^T$, but how to compare $(1,3)^T$ and $(3,1)^T$). This is always the case when vectors are compared in real spaces (see also Chankong and Haimes (1983b, pp. 64–67)).

Anyway, some of the objective vectors can be extracted for examination. Such vectors are those where none of the components can be improved without deterioration to at least one of the other components. Edgeworth (1987) presented this definition in 1881. However, the definition is usually called Pareto optimality after the French-Italian economist and sociologist Vilfredo Pareto, who in 1896 developed it further (see Pareto (1964, 1971)). However, in some connections, like in Stadler (1988b), the term *Edgeworth-Pareto optimality* is used for the above-mentioned reason. Koopmans was one of the first to employ in 1951 the concept of Pareto optimality in Koopmans (1971). A more formal definition of *Pareto optimality* is the following:

Definition 2.2.1. A decision vector $\mathbf{x}^* \in S$ is *Pareto optimal* if there does not exist another decision vector $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$ for at least one index j.

An objective vector $\mathbf{z}^* \in Z$ is Pareto optimal if there does not exist another objective vector $\mathbf{z} \in Z$ such that $z_i \leq z_i^*$ for all $i = 1, \ldots, k$ and $z_j < z_j^*$ for at least one index j; or equivalently, $\mathbf{z}^*$ is Pareto optimal if the decision vector corresponding to it is Pareto optimal.

In Figure 2.2.1, a feasible region $S \subset \mathbf{R}^3$ and its image, a feasible objective region $Z \subset \mathbf{R}^2$, are illustrated. The fat line contains all the Pareto optimal objective vectors. The vector $\mathbf{z}^*$ is an example of them.

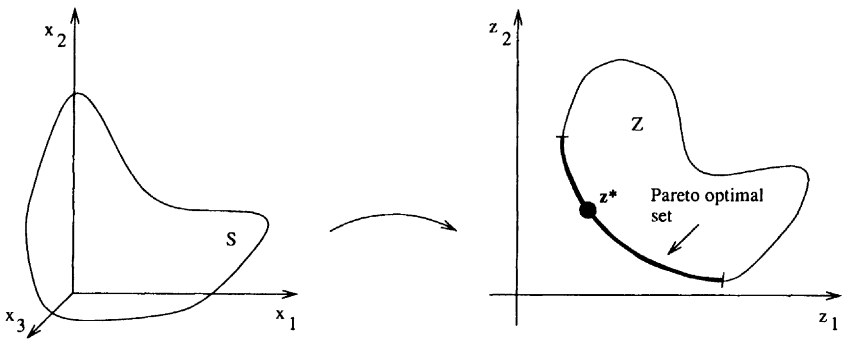

Figure 2.2.1. The sets S and Z and the Pareto optimal objective vectors.

There are usually a lot (infinite number) of Pareto optimal solutions. We can speak about a set of Pareto optimal solutions or a *Pareto optimal set*. This set can be nonconvex and nonconnected.

In addition to Pareto optimality, several other terms are sometimes used for the optimality concept described above. These terms are, for example, *noninferiority*, *efficiency* and *nondominance*. At variance with this practice, a more general meaning is given to efficiency later. In general, Pareto optimality is here used as a concept of optimality, unless stated otherwise.

Definition 2.2.1 introduces *global Pareto optimality*. Another important concept is local Pareto optimality.

Definition 2.2.2. A decision vector $\mathbf{x}^* \in S$ is *locally Pareto optimal* if there exists $\delta > 0$ such that $\mathbf{x}^*$ is Pareto optimal in $S \cap B(\mathbf{x}^*, \delta)$.

An objective vector $\mathbf{z}^* \in Z$ is locally Pareto optimal if the decision vector corresponding to it is locally Pareto optimal.

Naturally, any globally Pareto optimal solution is locally Pareto optimal. The converse is valid for convex multiobjective optimization problems. (For this result, see e.g., Censor (1977).)

Theorem 2.2.3. Let the multiobjective optimization problem be convex. Then every locally Pareto optimal solution is also globally Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be locally Pareto optimal. Thus there exist some $\delta > 0$ and a neighbourhood $B(\mathbf{x}^*, \delta)$ of $\mathbf{x}^*$ such that there is no $\mathbf{x} \in S \cap B(\mathbf{x}^*, \delta)$ for which $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and for at least one index j is $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$.

Let us assume that $\mathbf{x}^*$ is not globally Pareto optimal. In this case, there exists some other point $\mathbf{x}^\circ \in S$ such that

(2.2.1) $\quad f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*)$ for some j.

Let us define $\hat{\mathbf{x}} = \beta\mathbf{x}^\circ + (1 - \beta)\mathbf{x}^*$, where $0 < \beta < 1$ is selected such that $\hat{\mathbf{x}} \in B(\mathbf{x}^*, \delta)$. The convexity of S implies that $\hat{\mathbf{x}} \in S$.

By the convexity of the objective functions and employing (2.2.1), we obtain $f_i(\hat{\mathbf{x}}) \leq \beta f_i(\mathbf{x}^\circ) + (1 - \beta) f_i(\mathbf{x}^*) \leq \beta f_i(\mathbf{x}^*) + (1 - \beta) f_i(\mathbf{x}^*) = f_i(\mathbf{x}^*)$ for every $i = 1, \ldots, k$. Because $\mathbf{x}^*$ is locally Pareto optimal and $\hat{\mathbf{x}} \in B(\mathbf{x}^*, \delta)$, we must have $f_i(\hat{\mathbf{x}}) = f_i(\mathbf{x}^*)$ for all i.

Further, $f_i(\mathbf{x}^*) \leq \beta f_i(\mathbf{x}^\circ) + (1 - \beta) f_i(\mathbf{x}^*)$ for every $i = 1, \ldots, k$. Because $\beta > 0$, we can divide by it and obtain $f_i(\mathbf{x}^*) \leq f_i(\mathbf{x}^\circ)$ for all i. According to assumption (2.2.1), we have $f_j(\mathbf{x}^*) > f_j(\mathbf{x}^\circ)$ for some j. Here we have a contradiction. Thus, $\mathbf{x}^*$ is globally Pareto optimal. □

We can establish the above-mentioned result with somewhat weaker assumptions. It is sufficient to assume that all the objective functions are quasiconvex and strictly quasiconvex. This result has been treated, for example,

in Ruíz-Canales and Rufián-Lizana (1995). These assumptions can be further relaxed according to Luc and Schaible (1997).

Theorem 2.2.4. Let the multiobjective optimization problem have a convex feasible region and quasiconvex objective functions with at least one strictly quasiconvex objective function. Then every locally Pareto optimal solution is also globally Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be locally Pareto optimal. Thus there exist some $\delta > 0$ and a neighbourhood $B(\mathbf{x}^*, \delta)$ of $\mathbf{x}^*$ such that there is no $\mathbf{x} \in S \cap B(\mathbf{x}^*, \delta)$ for which $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and for at least one index j is $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$.

Let us assume that $\mathbf{x}^*$ is not globally Pareto optimal. In this case, there exists some other point $\mathbf{x}^\circ \in S$ such that

$$f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*) \text{ for all } i = 1, \ldots, k \text{ and } f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*) \text{ for some } j. \tag{2.2.2}$$

Let us define $\hat{\mathbf{x}} = \beta\mathbf{x}^\circ + (1 - \beta)\mathbf{x}^*$, where $0 < \beta < 1$ is selected such that $\hat{\mathbf{x}} \in B(\mathbf{x}^*, \delta)$. The convexity of S implies that $\hat{\mathbf{x}} \in S$.

Employing (2.2.2) and by the quasiconvexity of the objective functions, respectively, for each index i such that $f_i(\mathbf{x}^\circ) = f_i(\mathbf{x}^*)$, we obtain

$$f_i(\hat{\mathbf{x}}) \leq \max[f_i(\mathbf{x}^\circ), f_i(\mathbf{x}^*)] = f_i(\mathbf{x}^*),$$

and for each index j such that $f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*)$, we have

$$f_j(\hat{\mathbf{x}}) \leq \max[f_j(\mathbf{x}^\circ), f_j(\mathbf{x}^*)] = f_j(\mathbf{x}^*).$$

Because at least one of the objective functions is strictly quasiconvex, at least one of the inequalities above is strict. Here we have a contradiction with the local Pareto optimality of $\mathbf{x}^*$. Thus, $\mathbf{x}^*$ is globally Pareto optimal. □

For the sake of brevity, we shall usually speak only about Pareto optimality in the sequel. In practice, however, we only have locally Pareto optimal solutions computationally available, unless some additional requirement, such as convexity, is fulfilled.

Usually, we are interested in Pareto optimal solutions and can forget the other feasible solutions. Exceptions to this practice are problems where one of the objective functions is an approximation of an unknown function or there are underlying unexpressed objective functions involved. Then, the real Pareto optimal set is unknown.

According to the definition of Pareto optimality, moving from one Pareto optimal solution to another necessitates trading off. This is one of the basic concepts in multiobjective optimization. Let us, however, mention that the idea of trading off can be called into question, as suggested, for example, in Zeleny (1997). It is not perhaps always necessary to trade off in order to attain improved results. One can argue that it has been possible to produce things

both at lower cost and with higher quality. Changing the way of approaching the problem and its formulation may produce better results than simply trading off in the old formulation. (This can also be regarded as an example of expanding habitual domains, to be introduced in Section 2.3.) Zeleny goes so far as to claiming that trade-offs are properties of inadequately designed systems. For that reason one can claim that we should aim at designing systems better.

2.3. Decision Maker

Mathematically, every Pareto optimal point is an equally acceptable solution of the multiobjective optimization problem. However, it is generally desirable to obtain one point as a solution. Selecting one out of the set of Pareto optimal solutions calls for information that is not contained in the objective functions. This is why – compared to single objective optimization – a new element is added in multiobjective optimization.

We need a decision maker to make the selection. The *decision maker* is a person (or a group of persons) who is supposed to have better insight into the problem and who can express preference relations between different solutions. Usually, the decision maker is responsible for the final solution.

Solving a multiobjective optimization problem calls for the co-operation of the decision maker and an analyst. By an *analyst* we here mean a person or a computer program responsible for the mathematical side of the solution process. The analyst generates information for the decision maker to consider and the solution is selected according to the preferences of the decision maker.

It is assumed in the following that we have a single decision maker or a unanimous group of decision makers. Generally, group decision making is a world of its own. It calls for negotiations and specific methods when searching for compromises between different interest groups (see, for example, Hwang and Lin (1987) and Yu (1973)).

In Part II, solution methods are classified according to the role of the decision maker in the solution process. In some methods, various assumptions are made concerning the preference structure and behaviour of the decision maker. Note that assuming a single decision maker does not exclude the possibility that there may be others involved influencing the decision maker (as stressed in Zionts (1997a, b)).

During solution processes, various kinds of information are solicited from the decision maker. Such items of information may include, for example, desirable or acceptable levels in the values of the objective functions. These objective values (whether feasible or not) are of special interest and importance to the decision maker.

Definition 2.3.1. Objective function values that are satisfactory or desirable to the decision maker are called *aspiration levels* and denoted by $\bar{z}_i$, $i = 1, \ldots, k$. The vector $\bar{\mathbf{z}} \in \mathbf{R}^k$, consisting of aspiration levels, is called a *reference point*.

By solving a multiobjective optimization problem we here mean finding a feasible decision vector such that it is Pareto optimal and satisfies the needs and the requirements of the decision maker. Assuming such a solution exists, it is called a *final solution*. However, as stressed in Zionts (1997a, b), it may be difficult for the decision maker to distinguish between good and optimal solutions in real problems. If this is the case, the emphasis should be on finding good solutions (and sometimes, only, on finding solutions).

We do not focus here on the problems of decision making, which is a research area of its own. Interesting topics in this area are, for instance, decision making with incomplete information, validity of the problem formulation and habitual domains. The first of these matter is treated, for example, in Weber (1987). Reasons for incomplete information include lack of knowledge, pressure of time, fear of commitment and matters related to the future.

We usually assume that decision makers are only interested in Pareto optimal points and the rest can be excluded. However, this is not the case if the problem has not been formulated well enough. As already emphasized, non-Pareto optimal solutions may be important if there are some unformulated or hidden objective functions in the mind of the decision maker or some of the objective functions are simply proxies of the objective functions proper (see, for example, Zionts (1997a, b)). In such cases, the Pareto optimal sets of the problem handled and the actual problem which should be solved, do not coincide. Here we assume the mathematical model to be accurate and static so that we can mainly concentrate on Pareto optimal solutions.

A *habitual domain* is defined in Yu (1991) as a set of ways of thinking, judging and responding, as well as the knowledge and experience on which they are based. Yu emphasizes that in order to make effective decisions it is important to expand and enrich the habitual domains of the decision makers. Several ways of carrying this out are presented in Yu (1991, 1995). Understanding, expanding and enriching the domains of thinking is also stressed, for example, in Yu (1994) and Yu and Liu (1997).

2.4. Ranges of the Pareto Optimal Set

Let us for a while investigate the ranges of the set of Pareto optimal solutions. We assume that the objective functions are bounded over the feasible region S.

2.4.1. Ideal Objective Vector

An objective vector minimizing each of the objective functions is called an ideal (or perfect) objective vector.

Definition 2.4.1. The components $z_i^\star$ of the *ideal objective vector* $\mathbf{z}^\star \in \mathbf{R}^k$ are obtained by minimizing each of the objective functions individually subject to the constraints, that is, by solving

$$\begin{aligned} &\text{minimize} && f_i(\mathbf{x}) \\ &\text{subject to} && \mathbf{x} \in S, \end{aligned}$$

for $i = 1, \ldots, k$.

It is obvious that if the ideal objective vector were feasible (that is, $\mathbf{z}^\star \in Z$), it would be the solution of the multiobjective optimization problem (and the Pareto optimal set would be reduced to it). This is not possible in general since there is some conflict among the objectives. Even though the ideal objective vector is not attainable, it can be considered a reference point, something to go for. From the ideal objective vector we obtain the lower bounds of the Pareto optimal set for each objective function.

Note that in practice some caution is in order with nonconvex problems. The definition of the ideal objective vector assumes that we know the global minima of the individual objective functions. Guaranteeing global optimality in numerical calculations is not that simple. This must be kept in mind with practical problems. Properties of ideal objective vectors, for example, their uniqueness, are treated in Skulimowski (1992).

Sometimes we also need a vector that is strictly better than, in other words, strictly dominates, every Pareto optimal solution.

Definition 2.4.2. A *utopian objective vector* $\mathbf{z}^{\star\star} \in \mathbf{R}^k$ is an infeasible objective vector whose components are formed by

$$z_i^{\star\star} = z_i^\star - \varepsilon_i$$

for all $i = 1, \ldots, k$, where $z_i^\star$ is a component of the ideal objective vector and $\varepsilon_i > 0$ is a relatively small but computationally significant scalar.

2.4.2. Nadir Objective Vector

The upper bounds of the Pareto optimal set, that is, the components of a *nadir objective vector* (or imperfect objective vector) $\mathbf{z}^{\text{nad}}$, are much more difficult to obtain. However, they can be estimated from a payoff table.

A *payoff table* is formed by using the decision vectors obtained when calculating the ideal objective vector. Row i of the payoff table displays the values of all the objective functions calculated at the point where f_i obtained its minimal value. Hence, $z_i^\star$ is at the main diagonal of the table. The maximal value of the column i in the payoff table can be selected as an estimate of the upper bound of the objective f_i for $i = 1, \ldots, k$ over the Pareto optimal set.

The black points in Figure 2.4.1 represent ideal objective vectors, and the grey ones are nadir objective vectors. The nadir objective vector may be feasible

or not, as illustrated in Figure 2.4.1. The Pareto optimal set is represented by the bold lines.

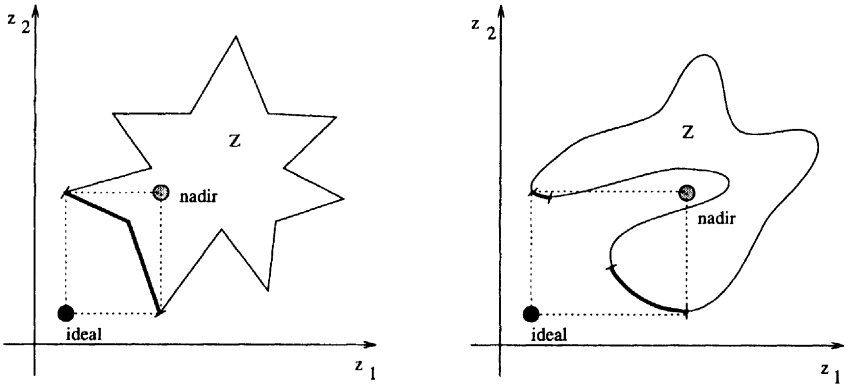

Figure 2.4.1. Ideal objective vectors and nadir objective vectors.

Note that the objective vectors in the rows of the payoff table are Pareto optimal if they are unique. In other words, if the individual objective functions have alternative optima, the obtained objective vector may not be Pareto optimal. This fact can weaken the approach and it can happen in linear as well as in nonlinear problems.

It is important to note that the estimates based on the payoff table are not necessarily equal to the real components of the nadir objective vector as demonstrated, for example, in Korhonen et al. (1997) and Weistroffer (1985). Instead of being correct, the nadir objective value approximate may be either far too low or too high.

The difference between the complete Pareto optimal set and the subset of the Pareto optimal set bounded by the ideal objective vector and the upper bounds obtained from the payoff table in linear cases is explored in Reeves and Reid (1988). It is proposed that relaxing (i.e., increasing) the approximated upper bounds by a relatively small tolerance should improve the approximation, although it is ad hoc in nature. However, small tolerances may not necessarily help because the error between the correct and the approximated nadir objective value may be significant.

For nonlinear problems, there is no constructive method for calculating the nadir objective vector. That is why we here mention some treatments for MOLP problems. Isermann and Steuer (1988) include an examination of how many of the Pareto optimal extreme solutions of some MOLP problems are above the upper bounds obtained from the payoff table. Three methods for determining the exact nadir objective vector in a linear case are also suggested. None of them is especially economical computationally. In Dessouky et al. (1986), three heuristics are presented for calculating the nadir objective vector when the

problem is linear. A heuristic for MOLP problems is also described in Korhonen et al. (1997). It is demonstrated how much better are the approximations the heuristic can provide. Heuristics are usually able to improve the approximations obtained from the payoff table even though they may not always find the correct nadir objective values. Heuristics are often computationally much less demanding than exact procedures.

Nonetheless, the payoff table may be used as a rough estimate as long as its robustness is kept in mind. Because of the above-described difficulty of calculating the actual nadir objective vector, we shall usually refer to the approximate nadir objective vector as $\mathbf{z}^{\text{nad}}$.

2.4.3. Related Topics

In many occasions it is advisable to rescale, that is, normalize the objective functions so that their objective values are of approximately the same magnitude. If the ideal objective vector and a good enough approximation to the nadir objective vector are known, we can replace each objective function $f_i(\mathbf{x})$ $(i = 1, \ldots, k)$ by the function

$$\frac{f_i(\mathbf{x}) - z_i^\star}{z_i^{\text{nad}} - z_i^\star}.$$

In this case, the range of each new objective function is $[0, 1]$.

Another related possibility is to use a *range equalization factor*, as suggested in Steuer (1986). The range R_i of each objective function is first estimated by the difference between the (possibly approximated) nadir objective vector and the ideal objective vector. Then, constants

$$K_i = \frac{1}{R_i} \frac{1}{\sum_{j=1}^{k} \frac{1}{R_j}}$$

are defined for every $i = 1, \ldots, k$, and finally each objective function is multiplied by K_i.

A simple alternative for normalizing the objective function values is to divide each objective function by its (nonzero) ideal objective value. This has been suggested, for example, in Osyczka (1984, 1992). This is not as exact as the previous methods but does not necessitate information about the nadir objective vector.

It is usually advisable to use normalized objective values only in calculations and to display restored objective values in the original scales to the decision maker. In this way the different scales do not confuse computation and significant objective values are offered to the decision maker.

It is possible that (some) objective functions are unbounded, for instance, from below. In this case some caution is in order. In multiobjective optimization problems this does not necessarily mean that the problem is formulated

incorrectly. There may still exist Pareto optimal solutions. However, if, for instance, some component of the ideal objective vector is unbounded and it is replaced by a small but finite number, methods utilizing the ideal objective vector may not be able to overcome the replacement.

Finally, let us look at some examples of the problem of optimizing a function over the Pareto optimal set of a multiobjective optimization problem. This is a more general problem than just looking for the ranges of the Pareto optimal set. In Benson and Sayin (1994), the authors deal with the maximization of a linear function over the Pareto optimal set of an MOLP problem. A general function is minimized over the Pareto optimal set of an MOLP problem in Dauer and Fosnaugh (1995), and a convex function is optimized over the Pareto optimal set of linear objective functions and a convex feasible region by duality techniques in Thach et al. (1996). Maximization of a function over the Pareto optimal set is also considered in Horst and Thoai (1997).

2.5. Weak Pareto Optimality

In addition to Pareto optimality, other related concepts are widely used. These are weak and proper Pareto optimality. The relationship between these concepts is that the properly Pareto optimal set is a subset of the Pareto optimal set which is a subset of the weakly Pareto optimal set.

A vector is weakly Pareto optimal if there does not exist any other vector for which all the components are better. More formally it means the following:

Definition 2.5.1. A decision vector $\mathbf{x}^* \in S$ is *weakly Pareto optimal* if there does not exist another decision vector $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$.

An objective vector $\mathbf{z}^* \in Z$ is weakly Pareto optimal if there does not exist another objective vector $\mathbf{z} \in Z$ such that $z_i < z_i^*$ for all $i = 1, \ldots, k$; or equivalently, if the decision vector corresponding to it is weakly Pareto optimal.

The bold line in Figure 2.5.1 represents the set of weakly Pareto optimal objective vectors. The fact that the Pareto optimal set is a subset of the weakly Pareto optimal set can also be seen in the figure. The Pareto optimal objective vectors are situated along the line between the dots.

Similarly to Pareto optimality, local weak Pareto optimality can be defined in addition to the global weak Pareto optimality of Definition 2.5.1. It must still be kept in mind that usually only locally weakly Pareto optimal solutions are computationally available. Nevertheless, for the sake of brevity, we shall usually refer only to weak Pareto optimality.

Let us state as a curiosity that if the feasible region is convex and the objective functions are quasiconvex with at least one strictly quasiconvex function, the set of locally Pareto optimal solutions is a subset of the set of weakly Pareto

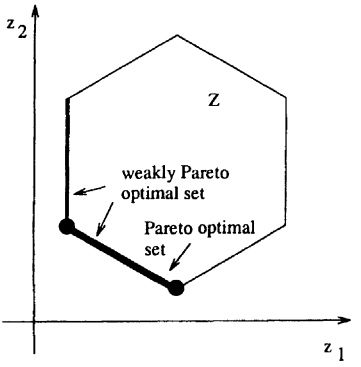

Figure 2.5.1. Weakly Pareto optimal vectors.

optimal solutions. This result is an immediate corollary of Theorem 2.2.4, where we proved that under the above-mentioned assumptions all the locally Pareto optimal solutions are also globally Pareto optimal.

The connectedness of the sets of Pareto optimal and weakly Pareto optimal solutions has not been widely treated. Yet, this is an important feature because it is often useful to know how well one can move continuously from one (weakly) Pareto optimal solution to another.

The Pareto optimal set of an MOLP problem is proved to be connected in Steuer (1986, pp. 158, 220). It is stated in Warburton (1983), that the Pareto optimal set is connected in convex multiobjective optimization problems. In addition, Warburton shows that if the feasible region is convex and compact and the objective functions are quasiconvex, then the set of weakly Pareto optimal solutions is connected. The connectedness of the Pareto optimal set is guaranteed for a certain subclass of quasiconvex functions. A noncompact case is also studied in Warburton (1983).

The structure, including connectedness, of the sets of weakly, properly or Pareto optimal solutions for nonconvex problems with two objective functions is investigated in Tenhuisen and Wiecek (1996). A review of connectedness results for Pareto optimality is given in Benoist (1998). Benoist also proves that the Pareto optimal set is connected for continuous, strictly quasiconvex objective functions (when transformed for minimization problems) defined on a convex and compact set.

Although weakly Pareto optimal solutions are important for theoretical considerations, they are not always useful in practice, because of the large size of the weakly Pareto optimal set. However, they are often relevant from a technical point of view because they are sometimes easier to generate than Pareto optimal points. A more restrictive concept than Pareto optimality is proper Pareto optimality (to be defined in Section 2.9).

2.6. Value Function

It is often assumed that the decision maker makes decisions on the basis of an underlying function of some kind. This function is called a value function.

Definition 2.6.1. A function $U\colon \mathbf{R}^k \to \mathbf{R}$ representing the preferences of the decision maker among the objective vectors is called a *value function*.

Let $\mathbf{z}^1$ and $\mathbf{z}^2 \in Z$ be two different objective vectors. If $U(\mathbf{z}^1) > U(\mathbf{z}^2)$, then the decision maker prefers $\mathbf{z}^1$ to $\mathbf{z}^2$. If $U(\mathbf{z}^1) = U(\mathbf{z}^2)$, then the decision maker finds the objective vectors equally desirable, that is, they are *indifferent*.

It must be pointed out that the value function is totally a decision maker-dependent concept. Different decision makers may have different value functions for the same problem.

Sometimes the term *utility function* is used instead of the value function. Here we follow the common way of referring to value functions in deterministic problems. The term utility function is reserved for stochastic problems (not to be handled here). See Keeney and Raiffa (1976) for a more extended discussion of both terms.

If we had at our disposal the mathematical expression of the decision maker's value function, it would be easy to solve the multiobjective optimization problem. The value function would simply be maximized by some method of single objective optimization. The value function would offer a total (complete) ordering of the objective vectors. However, there are several reasons why this seemingly easy way is not generally used in practice. The most important reason is that it is extremely difficult, if not impossible, for a decision maker to specify mathematically the function behind her or his preferences. Secondly, even if the function were known, it could be difficult to optimize because of its possible complicated nature. An example of such situations is the nonconcavity of the value function. In this case, only a local maximum may be found instead of the global one. In addition, as pointed out in Steuer and Gardiner (1991), it is not necessarily all to the good that optimizing the value function results in a single solution. After specifying the value function, the decision maker may have doubts about its validity. This is why (s)he may want to explore different alternatives before selecting the final solution.

One more thing to keep in mind about value functions is that their existence is not necessarily guaranteed. At least it may be restricting to assume that a fixed and stable function can explain the behaviour and the preferences of the decision maker.

Even though value functions are seldom explicitly used in solving multiobjective optimization problems, they are very important in the development of solution methods and as a theoretical background. In many multiobjective optimization methods, the value function is assumed to be known implicitly and the decision maker is assumed to make selections on this basis. In several

methods, convergence results are obtained by making certain assumptions, for example, quasiconcavity about the implicit value function. In all, we can say that value functions are usually more important to the analyst than to the decision maker (see Zionts (1997a, b)).

Generally, the value function is assumed to be strongly decreasing. This means that the preference of the decision maker will increase if the value of an objective function decreases while all the other objective values remain unchanged (i.e., less is preferred to more). This assumption is justified by Rosenthal (1985), who stresses that "Clearly, under the monotonicity assumption a rational decision maker would never deliberately select a dominated point. This is probably the only important statement in multiobjective optimization that can be made without the possibility of generating some disagreement."

However, there are exceptions to this situation. Rosenthal mentions as an (maximization) example the deer population, where more deer are usually preferred to fewer for aesthetic and recreational reasons, but not in the case when the deer population is large enough to remove all the forest undergrowth.

The following theorem presents an important result concerning the solutions of strongly decreasing value functions.

Theorem 2.6.2. Let the value function $U\colon \mathbf{R}^k \to \mathbf{R}$ be strongly decreasing. Let U attain its maximum at $\mathbf{z}^* \in Z$. Then $\mathbf{z}^*$ is Pareto optimal.

Proof. Let $\mathbf{z}^* \in Z$ be a maximal solution of a strongly decreasing value function U. Let us assume that $\mathbf{z}^*$ is not Pareto optimal. Then there exists an objective vector $\mathbf{z} \in Z$ such that $z_i \leq z_i^*$ for all $i = 1, \ldots, k$ and $z_j < z_j^*$ for at least one index j. Because U is strongly decreasing, we have $U(\mathbf{z}) > U(\mathbf{z}^*)$. Thus U does not attain its maximum at $\mathbf{z}^*$. This contradiction implies that $\mathbf{z}^*$ is Pareto optimal. □

Different properties and forms of value functions are widely treated in Hemming (1978). Some references handling the existence of value functions are listed in Stadler (1979) where different value functions are also presented.

The way a final solution was earlier defined means that a solution is final if it maximizes the decision maker's value function. Sometimes another concept, that of the satisficing solution, is distinguished.

Satisficing solutions are connected with so-called satisficing decision making. *Satisficing decision making* means that the decision maker does not intend to maximize any general value function but tries to achieve certain aspirations. A solution which satisfies all the aspirations of the decision maker is called a *satisficing solution*. In the most extreme case, one can define a solution to be satisficing independent of whether it is Pareto optimal or not. Here we, however, always assume that a satisficing solution is Pareto optimal or at least weakly Pareto optimal.

It is important to realize that regardless of the existence of an underlying value function, a general assumption still is that that less is preferred to more by the decision maker, that is, lower objective function values are preferred to higher. This assumption is usually made even in methods not involving value functions in any way. Thus, assuming that less is preferred to more is a more general assumption than assuming a strongly decreasing value function.

2.7. Efficiency

It is possible to define optimality in a multiobjective context in more general ways than by Pareto or weak Pareto optimality. Let us have a pointed convex cone D defined in $\mathbf{R}^k$. This cone D is called an *ordering cone* and it is used to induce a partial ordering on Z. Let us have two objective vectors, $\mathbf{z}^1$ and $\mathbf{z}^2 \in Z$. An objective vector $\mathbf{z}^1$ *dominates* $\mathbf{z}^2$, denoted by $\mathbf{z}^1 \leq_D \mathbf{z}^2$, if $\mathbf{z}^2 - \mathbf{z}^1 \in D$ and $\mathbf{z}^1 \neq \mathbf{z}^2$, that is, $\mathbf{z}^2 - \mathbf{z}^1 \in D \setminus \{\mathbf{0}\}$. The same can also be written as $\mathbf{z}^2 \in \mathbf{z}^1 + D$ and $\mathbf{z}^1 \neq \mathbf{z}^2$, that is, $\mathbf{z}^2 \in \mathbf{z}^1 + D \setminus \{\mathbf{0}\}$ as illustrated in Figure 2.7.1.

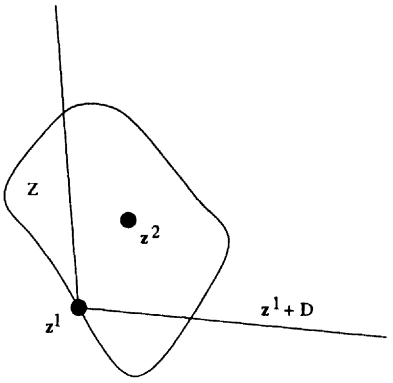

Figure 2.7.1. Domination induced by a cone D.

We can now present a definition of optimality based on domination, which is an alternative to the definitions previously given. When an ordering cone is used in defining optimality, then the term efficiency will be used in what follows.

Definition 2.7.1. Let D be a pointed convex cone. A decision vector $\mathbf{x}^* \in S$ is *efficient* (with respect to D) if there does not exist another decision vector $\mathbf{x} \in S$ such that $\mathbf{f}(\mathbf{x}) \leq_D \mathbf{f}(\mathbf{x}^*)$.

An objective vector $\mathbf{z}^* \in Z$ is efficient if there does not exist another objective vector $\mathbf{z} \in Z$ such that $\mathbf{z} \leq_D \mathbf{z}^*$.

This definition means that a vector is efficient (nondominated) if it is not dominated by any other feasible vector. The definition above can be formulated in many ways. If we substitute $\leq_D$ for its definition, we have the condition in the form $\mathbf{0} \neq \mathbf{z}^* - \mathbf{z} \in D$ or $\mathbf{z}^* - \mathbf{z} \in D \setminus \{\mathbf{0}\}$ (see Corley (1980)).

Other equivalent formulations are, for instance, $\mathbf{z}^* \in Z$ is efficient if $(Z - \mathbf{z}^*) \cap (-D) = \{\mathbf{0}\}$ (see Pascoletti and Serafini (1984) and Weidner (1988)), if $(\mathbf{z}^* - D \setminus \{\mathbf{0}\}) \cap Z = \emptyset$ (see Tapia and Murtagh (1989) and Wierzbicki (1986b)) or if $(\mathbf{z}^* - D) \cap Z = \mathbf{z}^*$ (see Chen (1984) and Jahn (1987)).

Let us give an alternative formulation to Definition 2.7.1 using one of the equivalent representations.

Definition 2.7.2. Let D be a pointed convex cone. A decision vector $\mathbf{x}^* \in S$ is *efficient* (with respect to D) if there does not exist another decision vector $\mathbf{x} \in S$ such that $\mathbf{f}(\mathbf{x}^*) \in \mathbf{f}(\mathbf{x}) + D \setminus \{\mathbf{0}\}$, that is, $(\mathbf{f}(\mathbf{x}^*) - D \setminus \{\mathbf{0}\}) \cap Z = \emptyset$.

An objective vector $\mathbf{z}^* \in Z$ is efficient if there does not exist another objective vector $\mathbf{z} \in Z$ such that $\mathbf{z}^* \in \mathbf{z} + D \setminus \{\mathbf{0}\}$, that is, $(\mathbf{z}^* - D \setminus \{\mathbf{0}\}) \cap Z = \emptyset$.

Different notions of efficiency are collected in Ester and Tröltzsch (1986). They provide several auxiliary problems in the interests of obtaining efficient solutions.

Remark 2.7.3. The above definitions are equivalent to Pareto optimality if $D = \mathbf{R}^k_+$ (see Figure 2.7.2).

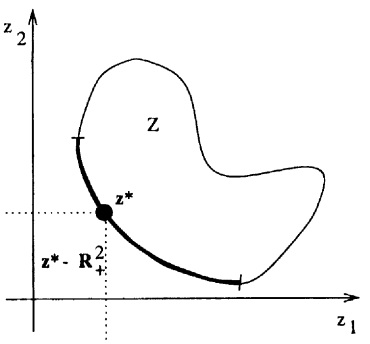

Figure 2.7.2. Pareto optimality with the help of cone $\mathbf{R}^k_+$.

When Pareto optimality or efficiency is defined with the help of ordering cones, it is trivial to verify that Pareto optimal or efficient objective vectors always lie on the boundary of the feasible objective region Z.

Instead of a cone D, which is constant for every objective vector, we can use a point-to-set map D from Z into $\mathbf{R}^k$ to represent the domination structure. In this case domination is dependent on the current objective vector. For details

of ordering cones, see Sawaragi et al. (1985, pp. 25–31) and Yu (1974, 1985, pp. 163–209).

Theorem 2.6.2 gives a relationship between Pareto optimal solutions and value functions. Relations can also be established between efficient solutions and value functions. To give an idea of them, let us consider a pseudoconcave value function U. According to pseudoconcavity whenever $\nabla U(\mathbf{z}^1)^T(\mathbf{z}^2 - \mathbf{z}^1) \leq 0$, we have $U(\mathbf{z}^2) \leq U(\mathbf{z}^1)$. We can now define an ordering cone as a map $D(\mathbf{z}) = \{\mathbf{d} \in \mathbf{R}^k \mid \nabla U(\mathbf{z})^T\mathbf{d} \leq 0\}$. This ordering cone can be used to determine efficient solutions. Note that if we have a value function, we can derive its domination structure, but not generally vice versa. See Yu (1974) for an example.

Weakly efficient decision and objective vectors can be defined in a corresponding fashion to efficient ones. If the set Z of objective vectors is ordered by an ordering cone D, weakly efficient vectors may be characterized in the following way (see Jahn (1987) and Wierzbicki (1986b)):

Definition 2.7.4. Let D be a pointed convex cone. A decision vector $\mathbf{x}^* \in S$ is *weakly efficient* (with respect to D) if there does not exist another decision vector $\mathbf{x} \in S$ such that $\mathbf{f}(\mathbf{x}^*) \in \mathbf{f}(\mathbf{x}) + \operatorname{int} D$, that is, $(\mathbf{f}(\mathbf{x}^*) - \operatorname{int} D) \cap Z = \emptyset$.

An objective vector $\mathbf{z}^* \in Z$ is weakly efficient if there does not exist another objective vector $\mathbf{z} \in Z$ such that $\mathbf{z}^* \in \mathbf{z} + \operatorname{int} D$, that is, $(\mathbf{z}^* - \operatorname{int} D) \cap Z = \emptyset$.

An alternative formulation is that an objective vector $\mathbf{z}^* \in Z$ is weakly efficient if $(Z - \mathbf{z}^*) \cap (-\operatorname{int} D) = \emptyset$ (see Sawaragi et al. (1985, pp. 33–34)).

Connectedness of the sets of weakly efficient and efficient points is studied in Helbig (1990) whereas Luc (1989, pp. 148–154) treats particularly weakly efficient sets in convex problems where the objective functions are quasiconvex. In addition, connectedness results for efficient points in multiobjective combinatorial problems are given in Ehrgott and Klamroth (1997).

In the following, we mostly settle for treating Pareto optimality. Some extensions related to efficiency are only mentioned in passing.

Thus far, we have defined Pareto and weak Pareto optimality and more general efficiency and weak efficiency. Proper Pareto optimality and proper efficiency are yet to be introduced. To clarify their practical meaning and for other further purposes we must first, however, define trade-offs and marginal rates of substitution.

2.8. From One Solution to Another

Trade-offs and marginal rates of substitution are related to changes in the objective values when we move from one solution to another. Trade-offs are defined mathematically whereas marginal rates of substitution depend on the decision maker.

2.8.1. Trade-Offs

We have several concepts involved in trading off. A *trade-off* reflects the ratio of change in the values of the objective functions concerning the increment of one objective function that occurs when the value of some other objective function decreases. In the following definitions we have $i,\ j = 1, \ldots, k,\ i \neq j$.

Definition 2.8.1. (From Chankong and Haimes (1983b)) Let $\mathbf{x}^1$ and $\mathbf{x}^2 \in S$ be two decision vectors and let $\mathbf{f}(\mathbf{x}^1)$ and $\mathbf{f}(\mathbf{x}^2)$ be the corresponding objective vectors, respectively. We denote the ratio of change between the functions f_i and f_j by

$$\Lambda_{ij} = \Lambda_{ij}(\mathbf{x}^1, \mathbf{x}^2) = \frac{f_i(\mathbf{x}^1) - f_i(\mathbf{x}^2)}{f_j(\mathbf{x}^1) - f_j(\mathbf{x}^2)},$$

where $f_j(\mathbf{x}^1) - f_j(\mathbf{x}^2) \neq 0$.

Now, Λ_{ij} is called a *partial trade-off*, involving f_i and f_j between $\mathbf{x}^1$ and $\mathbf{x}^2$ if $f_l(\mathbf{x}^1) = f_l(\mathbf{x}^2)$ for all $l = 1, \ldots, k,\ l \neq i, j$. If $f_l(\mathbf{x}^1) \neq f_l(\mathbf{x}^2)$ for at least one $l = 1, \ldots, k$, and $l \neq i, j$, then Λ_{ij} is called a *total trade-off*, involving f_i and f_j between $\mathbf{x}^1$ and $\mathbf{x}^2$.

Note that in the case of two objective functions there is no difference between partial and total trade-offs. If partial trade-offs are presented to the decision maker, (s)he can compare changes in two objective functions at a time. This is usually a more comfortable procedure than comparing several objectives. If the points $\mathbf{x}^1$ and $\mathbf{x}^2$ are Pareto optimal, then there always exist some objective functions f_i and f_j for which the trade-off is negative. A concept related to the trade-off is the trade-off rate.

Definition 2.8.2. (From Chankong and Haimes (1983b)) Let $\mathbf{x}^* \in S$ be a decision vector and let $\mathbf{d}^*$ be a feasible direction emanating from $\mathbf{x}^*$. The *total trade-off rate* at $\mathbf{x}^*$, involving f_i and f_j along the direction $\mathbf{d}^*$, is given by

$$\lambda_{ij} = \lambda_{ij}(\mathbf{x}^*, \mathbf{d}^*) = \lim_{\alpha \to 0^+} \Lambda_{ij}(\mathbf{x}^* + \alpha \mathbf{d}^*, \mathbf{x}^*).$$

If $\mathbf{d}^*$ is a feasible direction so that there exists $\bar{\alpha} > 0$ satisfying $f_l(\mathbf{x}^* + \alpha \mathbf{d}^*) = f_l(\mathbf{x}^*)$ for all $l = 1, \ldots, k,\ l \neq i, j$ and for all $0 \leq \alpha \leq \bar{\alpha}$, then the corresponding λ_{ij} is called a *partial trade-off rate*.

Remark 2.8.3. If the objective functions are continuously differentiable, then

$$\lambda_{ij} = \frac{\nabla f_i(\mathbf{x}^*)^T \mathbf{d}^*}{\nabla f_j(\mathbf{x}^*)^T \mathbf{d}^*},$$

where the denominator differs from zero.

For continuously differentiable objective functions we can alternatively give the following definition.

Definition 2.8.4. Let the objective functions be continuously differentiable at a decision vector $\mathbf{x}^* \in S$. Then a partial trade-off rate at $\mathbf{x}^*$, involving f_i and f_j, is given by

$$\lambda_{ij} = \lambda_{ij}(\mathbf{x}^*) = \frac{\partial f_i(\mathbf{x}^*)}{\partial f_j}.$$

Differing from the idea of the definitions above, a so-called global trade-off is defined in Kaliszewski and Michalowski (1995, 1997). A global trade-off involves two objective functions and one decision vector which does not have to be Pareto optimal. It is the largest pairwise trade-off of two objective functions for one decision vector. Let us consider $\mathbf{x}^* \in S$ and modify the definitions for minimization problems. We define a subset of the feasible decision vectors in the form

$$S_j^{>}(\mathbf{x}^*) = \{\mathbf{x} \in S \mid f_j(\mathbf{x}) > f_j(\mathbf{x}^*),\ f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*),\ \text{ for } i = 1, \ldots, k,\ i \neq j\}.$$

Now we can introduce global trade-offs.

Definition 2.8.5. (From Kaliszewski and Michalowski (1995, 1997)) Let $\mathbf{x}^* \in S$ be a decision vector. We denote a *global trade-off* between the functions f_i and f_j by

$$\Lambda_{ij}^G = \Lambda_{ij}^G(\mathbf{x}^*) = \sup_{\mathbf{x} \in S_j^{>}(\mathbf{x}^*)} \frac{f_i(\mathbf{x}^*) - f_i(\mathbf{x})}{f_j(\mathbf{x}) - f_j(\mathbf{x}^*)}.$$

If $S_j^{>}(\mathbf{x}^*) = \emptyset$, then $\Lambda_{ij}^G(\mathbf{x}^*) = -\infty$ for every $i = 1, \ldots, k$, $i \neq j$.

A generalized definition of trade-offs in terms of tangent cones, meaning feasible directions, in the objective space is presented in Henig and Buchanan (1994, 1997). These generalized trade-off directions can be used for calculating trade-off rates at every Pareto optimal point of a convex multiobjective optimization problem.

Note that trade-offs are defined mathematically and the decision maker cannot affect them. If we take into consideration the opinions of the decision maker, we can define indifference curves and marginal rates of substitution.

2.8.2. Marginal Rate of Substitution

It is said that two feasible solutions are situated on the same *indifference curve* (or isopreference curve) if the decision maker finds them equally desirable, that is, neither of them is preferred to the other one. This means that indifference curves are contours of the underlying value function. There may also be a 'wider' *indifference band*. In this case we do not have any well-defined boundary between preferences, but a band where indifference occurs. This concept is studied in Passy and Levanon (1984).

For any two solutions on the same indifference curve there is a trade-off involving a certain increment in the value of one objective function (f_j) that the decision maker is willing to tolerate in exchange for a certain amount of decrement in some other objective function (f_i) while the preferences of the two solutions remain the same. This is called the marginal rate of substitution. This kind of trading between different solutions is characteristic of multiobjective optimization problems when moving from one Pareto optimal solution to another. The marginal rate of substitution (sometimes also called indifference trade-off) is the negative of the slope of the tangent to the indifference curve at a certain point.

Definition 2.8.6. A *marginal rate of substitution* $m_{ij} = m_{ij}(\mathbf{x}^*)$ represents the preferences of the decision maker at a decision vector $\mathbf{x}^* \in S$. It is the amount of decrement in the value of the objective function f_i that compensates the decision maker for the one-unit increment in the value of the objective function f_j, while the values of all the other objectives remain unaltered.

Note that in the definition the starting and the resulting objective vectors lie on the same indifference curve and $i,\ j = 1, \ldots, k,\ i \neq j$.

It can be stated that the final solution of a multiobjective optimization problem is a Pareto optimal solution where the indifference curve is tangent to the Pareto optimal set. This tangency condition means finding an indifference curve intersecting the feasible objective region that is farthest to the southwest. This property is illustrated in Figure 2.8.1.

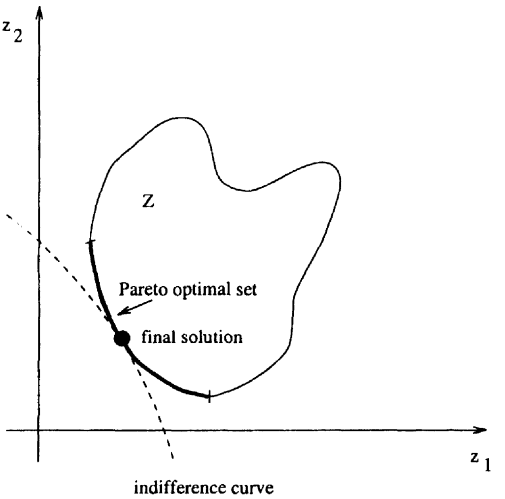

Figure 2.8.1. The final solution.

Remark 2.8.7. If the partial derivatives exist, then

$$m_{ij}(\mathbf{x}^*) = \frac{\partial U(\mathbf{f}(\mathbf{x}^*))}{\partial f_j} \Big/ \frac{\partial U(\mathbf{f}(\mathbf{x}^*))}{\partial f_i}.$$

If the Pareto optimal set is smooth (that is, at every Pareto optimal point there exists a unique tangent), we have the following result. When one examines the definition of a trade-off rate at some point, one sees that it is the slope of the tangent of the Pareto optimal set at that point. We can also define that when a Pareto optimal solution is a final solution, then the tangents of the indifference curve and the Pareto optimal set coincide at it, that is,

$$-m_{ij} = \lambda_{ij} \quad \text{for all} \quad i, j = 1, \ldots, k, \; i \neq j. \tag{2.8.1}$$

Thus, with the help of the negative of the marginal rate of substitution and the trade-off rate one can get a local linear approximation of the indifference curve and the Pareto optimal set, respectively.

Usually, one of the objective functions is selected as a *reference function* when trade-offs and marginal rates of substitution are treated. The trade-offs and the marginal rates of substitution are generated with respect to it. In the notations above, f_i is the reference function. When co-operating with decision makers, it is important to select the reference function in a meaningful way. An important criterion in the selection is, for example, that the reference function is in familiar units or that it is dominant.

2.9. Proper Pareto Optimality

Kuhn and Tucker were the first to note that some of the Pareto optimal solutions had undesirable properties (see Kuhn and Tucker (1951)). To avoid such properties, they introduced properly Pareto optimal solutions and suggested that Pareto optimal solutions be divided into properly and improperly Pareto optimal ones. The idea of properly Pareto optimal solutions is that unbounded trade-offs between objectives are not allowed. Practically, a properly Pareto optimal solution with very high or very low trade-offs does not essentially differ from a weakly Pareto optimal solution for a human decision maker.

There exist several definitions for proper Pareto optimality. The idea is easiest to understand from the following definition.

Definition 2.9.1. (From Geoffrion (1968)) A decision vector $\mathbf{x}^* \in S$ is *properly Pareto optimal (in the sense of Geoffrion)* if it is Pareto optimal and if there is some real number $M > 0$ such that for each f_i and each $\mathbf{x} \in S$ satisfying $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$, there exists at least one f_j such that $f_j(\mathbf{x}^*) < f_j(\mathbf{x})$ and

$$\frac{f_i(\mathbf{x}^*) - f_i(\mathbf{x})}{f_j(\mathbf{x}) - f_j(\mathbf{x}^*)} \leq M.$$

An objective vector $\mathbf{z}^* \in Z$ is properly Pareto optimal if the decision vector corresponding to it is properly Pareto optimal.

In other words, a solution is properly Pareto optimal if there is at least one pair of objectives for which a finite decrement in one objective is possible only at the expense of some reasonable increment in the other objective.

Geoffrion's definition can be generalized so that the upper bound is a positive function $M(\mathbf{x})$ instead of a constant (see Mishra (1996) and Mishra and Mukherjee (1995)). This leads to the definition of *conditional proper Pareto optimality*.

A method for obtaining all the properly Pareto optimal solutions satisfying prescribed marginal rates of substitution in the convex case is proposed in Geromel and Ferreira (1991). Upper estimates for properly Pareto optimal solutions are given as well.

Durier (1988) studies the relationships between Pareto optimal and properly Pareto optimal sets in a convex case. One of the results is that if the set of properly Pareto optimal solutions is closed, then the two sets are equal. A property called a locally flat surface, which guarantees the very same equality in convex and differentiable problems, is presented in Zhou et al. (1993).

Results concerning Pareto optimal and properly Pareto optimal solutions are collected in Gal (1986). In Chew and Choo (1984), it is proved that every Pareto optimal solution is also properly Pareto optimal for a nonlinear problem involving only pseudolinear functions (i.e., differentiable functions which are both pseudoconvex and pseudoconcave). The results of Chew and Choo can be considered special cases of more general results presented in Weir (1990). In Gulati and Islam (1990), it is shown that the preceding result can be generalized by assuming quasiconvexity of the active constraints (of the form $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$) with some regularity properties. Pseudolinearity is extended by defining semilocally pseudolinear functions in Kaul et al. (1988).

We shall present some results concerning the relationships between Pareto optimal, weakly and properly Pareto optimal solutions in the context of solution methods in Part II.

Next, we introduce ε-proper Pareto optimality, which is easy to illustrate graphically.

Definition 2.9.2. (From Wierzbicki (1980b)) A decision vector $\mathbf{x}^* \in S$ and the corresponding objective vector $\mathbf{z}^* \in Z$ are *ε-properly Pareto optimal* if

$$(\mathbf{z}^* - \mathbf{R}^k_\varepsilon \setminus \{\mathbf{0}\}) \cap Z = \emptyset,$$

where $\mathbf{R}^k_\varepsilon = \{\mathbf{z} \in \mathbf{R}^k \mid \operatorname{dist}(\mathbf{z}, \mathbf{R}^k_+) \leq \varepsilon \|\mathbf{z}\|\}$ or $\mathbf{R}^k_\varepsilon = \{\mathbf{z} \in \mathbf{R}^k \mid \max_{i=1,\ldots,k} z_i + \varepsilon \sum_{i=1}^k z_i \geq 0\}$ and $\varepsilon > 0$ is a predetermined scalar.

Note that this definition differs from that of Pareto optimality so that a larger set $\mathbf{R}_\varepsilon^k$ is used instead of the set $\mathbf{R}_+^k$. The set of ε-properly Pareto optimal solutions is depicted in Figure 2.9.1 and denoted by a bold line. The solutions are obtained by intersecting the feasible objective region with a blunt cone. The end points of the Pareto optimal set, $\mathbf{z}^1$ and $\mathbf{z}^2$, have also been marked to ease the comparison.

An alternative formulation of Definition 2.9.2 is that a decision vector $\mathbf{x}^* \in S$ and the corresponding $\mathbf{z}^* \in Z$ are ε-properly Pareto optimal if $(\mathbf{z}^* - \mathbf{R}_\varepsilon^k) \cap Z = \mathbf{z}^*$. (The definition can be generalized into proper efficiency by using a convex cone D such that $\mathbf{R}_+^k \subset \operatorname{int} D \cup \{\mathbf{0}\}$.)

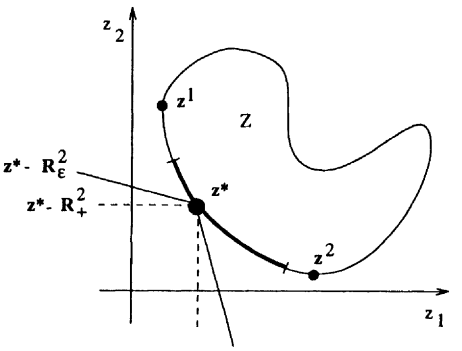

Figure 2.9.1. The set of ε-properly Pareto optimal solutions.

An interesting aspect of ε-properly Pareto optimal solutions is that the trade-offs are bounded by ε and $1/\varepsilon$ (see Wierzbicki (1986a, b)). We return to this concept in Section 3.5 of Part II.

Before we continue with the original definition of Kuhn and Tucker, we should mention briefly another way of decreasing the set of Pareto optimal solutions according to Liu (1996). There, $\mathbf{z}^* \in Z$ is called ϵ-Pareto optimal if $(\mathbf{z}^* - (\mathbf{R}_+^k + \boldsymbol{\epsilon}) \setminus \{\mathbf{0}\}) \cap Z = \emptyset$, where $\boldsymbol{\epsilon} \in \mathbf{R}_+^k$.

Let us for a while assume that the feasible region is defined with the help of inequality constraints. In other words, $S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}$. In addition, all the objective and the constraint functions are assumed to be continuously differentiable at every point $\mathbf{x} \in S$. Thus the next definition is not applicable to nondifferentiable multiobjective optimization problems.

Definition 2.9.3. (From Kuhn and Tucker (1951)) A decision vector $\mathbf{x}^* \in S$ is *properly Pareto optimal (in the sense of Kuhn and Tucker)* if it is Pareto optimal and if there does not exist any vector $\mathbf{d} \in \mathbf{R}^n$ such that

$$\nabla f_i(\mathbf{x}^*)^T \mathbf{d} \leq 0$$

for all $i = 1, \ldots, k$, for some j

$$\nabla f_j(\mathbf{x}^*)^T \mathbf{d} < 0,$$

and

$$\nabla g_l(\mathbf{x}^*)^T \mathbf{d} \leq 0$$

for all l satisfying $g_l(\mathbf{x}^*) = 0$, that is, for all active constraints at $\mathbf{x}^*$.

An objective vector $\mathbf{z}^* \in Z$ is properly Pareto optimal if the decision vector corresponding to it is properly Pareto optimal.

Kuhn and Tucker also derived necessary and sufficient conditions for proper Pareto optimality in Kuhn and Tucker (1951). Those conditions will be presented in the next section.

A comparison of the definitions of Kuhn and Tucker and Geoffrion is presented in Geoffrion (1968). For example, in convex cases the definition of Kuhn and Tucker implies the definition of Geoffrion. The reverse result is valid if the so-called Kuhn-Tucker constraint qualification (see Definition 3.1.3) is satisfied. The relationships of these two definitions are also treated, for example, in Sawaragi et al. (1985, pp. 42–46). Several practical examples are given in Tamura and Arai (1982) to illustrate the fact that properly Pareto optimal solutions according to the definitions of Kuhn and Tucker and Geoffrion (and one more definition by Klinger; see Klinger (1967)) are not necessarily consistent. Conditions under which (local) proper Pareto optimality in the sense of Kuhn and Tucker implies (local) proper Pareto optimality in the sense of Geoffrion are proved as well. More mathematical results concerning the properties and the relationships of the definitions of Kuhn and Tucker, Geoffrion and Klinger are given in White (1983a).

Borwein (1977) and Benson (1979a) have both defined proper efficiency when a closed, convex cone D is used as an ordering cone. Borwein's definition is based on tangent cones and Benson's on so-called projecting cones. Let us mention that proper efficiency according to Benson's definition implies proper efficiency in the sense of Borwein. (The reverse is valid in convex cases.) These two definitions are generalized in Henig (1982b) using convex ordering cones. The ordering cone D used in defining efficiency is utilized in the following.

Definition 2.9.4. (From Henig (1982b)) Let D be a pointed convex cone. A decision vector $\mathbf{x}^* \in S$ is *properly efficient (in the sense of Henig)* (with respect to D) if there does not exist another decision vector $\mathbf{x} \in S$ such that $\mathbf{f}(\mathbf{x}^*) \in \mathbf{f}(\mathbf{x}) + E \setminus \{\mathbf{0}\}$, that is, $(\mathbf{f}(\mathbf{x}^*) - E \setminus \{\mathbf{0}\}) \cap Z = \emptyset$ for some convex cone E such that $D \setminus \{\mathbf{0}\} \subset \operatorname{int} E$.

An objective vector $\mathbf{z}^* \in Z$ is properly efficient if there does not exist another objective vector $\mathbf{z} \in Z$ such that $\mathbf{z}^* \in \mathbf{z} + E \setminus \{\mathbf{0}\}$, in other words, $(\mathbf{z}^* - E \setminus \{\mathbf{0}\}) \cap Z = \emptyset$ with E as above.

The desirable property is valid also here: if a point is properly efficient, it is also efficient. Notice that Definition 2.9.4 is related to Definition 2.9.2 when we set $D = \mathbf{R}^k_+$.

As pointed out, different definitions of proper efficiency (and proper Pareto optimality) are not equivalent with each other but they have connections. The relationships between the definitions in the sense of Kuhn and Tucker, Geoffrion, Borwein, Benson and Henig are analysed in Sawaragi et al. (1985, pp. 39–44). For instance, Geoffrion's and Benson's definitions are equal when $D = \mathbf{R}^k_+$ (see also Benson (1983)). On the other hand, Definition 2.9.4 is equivalent to Benson's definition if the ordering cone D is closed and its closure is pointed. For further analysis we refer to Sawaragi et al. (1985, pp. 39–44).

In Henig (1982b), necessary and sufficient conditions for the existence of properly efficient solutions are given.

Let us finally mention that a new kind of proper efficiency, called super efficiency, is suggested in Borwein and Zhuang (1991, 1993).

In the following, proper Pareto optimality is understood in the sense of Geoffrion unless stated otherwise.

2.10. Pareto Optimality Tests with Existence Results

Let us have a look at how the Pareto optimality of feasible decision vectors can be tested. The procedures presented can also be used to find an initial Pareto optimal solution for (interactive) solution methods or to examine the existence of Pareto optimal and properly Pareto optimal solutions.

Specific results for MOLP problems are presented in Ecker and Kouada (1975). They are generalized for nonlinear problems with the help of duality theory in Wendell and Lee (1977). The treatment is based on an auxiliary problem

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{k} f_i(\mathbf{x}) \\ &\text{subject to} && f_i(\mathbf{x}) \leq f_i(\hat{\mathbf{x}}) \quad \text{for all} \quad i = 1, \ldots, k, \\ & && \mathbf{x} \in S, \end{aligned} \tag{2.10.1}$$

where $\hat{\mathbf{x}}$ is any vector in S. Let us denote the optimal objective function value by $\phi(\hat{\mathbf{x}})$.

Theorem 2.10.1. Let a decision vector $\mathbf{x}^* \in S$ be given. The vector $\mathbf{x}^*$ is Pareto optimal if and only if it is a solution of problem (2.10.1) so that $\phi(\mathbf{x}^*) = \sum_{i=1}^{k} f_i(\mathbf{x}^*)$.

On the other hand, let $\mathbf{x}^* \in S$ be a solution of problem (2.10.1). Then $\mathbf{x}^*$ is Pareto optimal and $f_i(\mathbf{x}^*) \leq f_i(\hat{\mathbf{x}})$ for all $i = 1, \ldots, k$.

Proof. See Wendell and Lee (1977).

Theorem 2.10.1 means that if problem (2.10.1) has an optimal solution for some $\hat{\mathbf{x}} \in S$, then either $\hat{\mathbf{x}}$ is Pareto optimal or the optimal solution of (2.10.1) is.

When studying the (primal) problem (2.10.1) and its dual, a *duality gap* is said to occur if the optimal value of the primal problem is not equivalent to the optimal value of the dual problem.

Theorem 2.10.2. Let a decision vector $\hat{\mathbf{x}} \in S$ be given and assume that $\phi(\hat{\mathbf{x}}) = -\infty$. Then some $\mathbf{x}^* \in S$ is Pareto optimal only if there is a duality gap between the primal (2.10.1) and its dual problem at $\mathbf{x}^*$. If such a gap exists, the optimal solution of (2.10.1) is Pareto optimal.

Proof. See Wendell and Lee (1977).

The significance of Theorem 2.10.2 is that precluding duality gaps the nonexistence of Pareto optimal points is characterized by the condition that $\phi(\hat{\mathbf{x}}) = -\infty$ for some $\hat{\mathbf{x}} \in S$. It can also be proved that if a multiobjective optimization problem is convex and if $\phi(\hat{\mathbf{x}}) = -\infty$ for some $\hat{\mathbf{x}} \in S$, then no properly Pareto optimal solutions exist. See the details in Wendell and Lee (1977).

Tests for Pareto optimality and the existence of Pareto optimal and properly Pareto optimal solutions are also investigated in Benson (1978). The results can be combined into the following theorem.

Theorem 2.10.3. Let a decision vector $\mathbf{x}^* \in S$ be given. Solve the problem

$$
\begin{aligned}
&\text{maximize} && \sum_{i=1}^{k} \varepsilon_i \\
&\text{subject to} && f_i(\mathbf{x}) + \varepsilon_i = f_i(\mathbf{x}^*) \quad \text{for all} \quad i = 1, \ldots, k, \\
& && \varepsilon_i \geq 0 \quad \text{for all} \quad i = 1, \ldots, k, \\
& && \mathbf{x} \in S,
\end{aligned}
\tag{2.10.2}
$$

where both $\mathbf{x} \in \mathbf{R}^n$ and $\boldsymbol{\varepsilon} \in \mathbf{R}_+^k$ are variables. Then the following results are valid.

(1) The vector $\mathbf{x}^*$ is Pareto optimal if and only if problem (2.10.2) has an optimal objective function value of zero.
(2) If problem (2.10.2) has a finite nonzero optimal objective function value obtained at a point $\hat{\mathbf{x}}$, then $\hat{\mathbf{x}}$ is Pareto optimal.
(3) If the multiobjective optimization problem is convex and if problem (2.10.2) does not have a finite optimal objective function value, then the set of properly Pareto optimal solutions is empty.

(4) If in addition to the conditions in (3), the set $\{\hat{\mathbf{z}} \in \mathbf{R}^k \mid \hat{\mathbf{z}} \leq \mathbf{f}(\mathbf{x})$ for some $\mathbf{x} \in S\}$ is closed, then the Pareto optimal set is empty.

Proof. See Benson (1978) or Chankong and Haimes (1983b, pp. 151–152).

Problem (2.10.2) is a popular way of checking Pareto optimality and of generating Pareto optimal solutions. However, sometimes equality constraints cause computational difficulties. Therefore it is useful to note that the equalities in (2.10.2) can be replaced with inequalities $f_i(\mathbf{x}) + \varepsilon_i \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ without affecting the generality of the results presented.

Two simple tests are suggested in Brosowski and da Silva (1994) for determining whether a given point is (locally) Pareto optimal or not. The tests are not based on any scalarizing functions but linear systems of equations. There are, however, several limitations. The objective functions are assumed to be continuously differentiable and their number has to be strictly larger than the number of variables. Further, no constraints can be included. Finally, the tests may also fail as demonstrated in Brosowski and da Silva (1994).

It is proved in Sawaragi et al. (1985, p. 59), that Pareto optimal solutions exist to multiobjective optimization problems where all the objective functions are lower semicontinuous (more general than continuity) and the feasible region is compact. Several ways of determining the Pareto optimality of a particular point in an MOLP problem are presented in Eiselt et al. (1987). They all apply to special situations. Further, the existence of Pareto optimal solutions when there is an infinite number of objective functions is considered in Alekseichik and Naumov (1981).

The existence of weakly Pareto optimal solutions in convex differentiable multiobjective optimization problems is treated in Deng (1998a). In addition, the compactness of the weakly Pareto optimal set is considered. The nonemptiness of the Pareto optimal and the weakly Pareto optimal sets in convex problems is also characterized in Deng (1998b).

As mentioned, auxiliary problems (2.10.1) and (2.10.2) can be used to produce Pareto optimal solutions, for example, from weakly Pareto optimal solutions. However, in some practical problems it is very expensive to carry out these additional optimizations. An alternative is suggested in Helbig (1991). If optimality is defined by an ordering cone, efficient solutions can be generated by perturbing this cone. In other words, using a method producing weakly efficient solutions with respect to the perturbed cone gives results that are efficient vis a vis the original problem.

The existence and the characterization of efficient solutions with respect to ordering cones are studied in Henig (1982a), and the existence of efficient solutions in linear spaces is treated in Borwein (1983). In addition, the existence of weakly and properly efficient (in the sense of Borwein) and efficient solutions in the presence of ordering cones is studied in Jahn (1986b). The existence of

efficient solutions is also treated in Cambini and Martein (1994) by introducing so-called quasi-D-bounded sets.

A phenomenon called *complete efficiency* occurs when every feasible decision vector of a multiobjective optimization problem is Pareto optimal. Tests are presented in Benson (1991) to check for complete efficiency in linear and nonlinear cases. A significant saving of computational efforts can be attained if the problem is tested for complete efficiency before it is solved. If the problem is completely efficient, no time, effort and special machinery for generating some or all of the Pareto optimal solutions is needed. Anyway, no solution algorithm exists which first checks for complete efficiency. The frequency of completely efficient problems among multiobjective optimization problems deserves further study. It may be more common than is generally thought, especially with special problem types, for example, when the feasible region S has no interior, as Benson points out. Transportation problems feature in this category. Complete efficiency is also treated in Weidner (1990).

One further area of research concerns the *domination property*. It refers to the situation where there always exists an efficient solution that is superior to any nonefficient solution, that is, for each $\mathbf{x} \in S$ and corresponding $\mathbf{z} \in Z$ there exists an efficient point $\mathbf{x}^*$ and corresponding $\mathbf{z}^*$ such that $\mathbf{z} - \mathbf{z}^* \in D$, where D is the ordering cone. Validity conditions for the domination property are examined in Benson (1983). The results of Benson are corrected and necessary and sufficient conditions for the domination property to hold are supplied in Luc (1984a). The domination property and its sufficient conditions are also treated in Henig (1986). Further, it is demonstrated that the existence of an efficient solution, the existence of a properly efficient solution, and the domination property are equivalent in solving convex problems. The domination property in infinite-dimensional spaces and for the sum of two sets is handled in Luc (1990).

The last concept to be mentioned here is the *redundancy* of objective functions. In MOLP cases this can be understood as linear dependency. In other words, an objective function is redundant if it does not affect the Pareto optimal set (see Gal and Leberling (1977)). This is not necessarily valid for nonlinear problems or in connection with interactive methods. For both of these, it is important to define redundancy on the basis of conflict between the objectives, which is why in Agrell (1997), an objective function is defined as redundant if it is not in conflict with any other objective function. Agrell suggests a probabilistic Monte-Carlo simulation-based redundancy test for nonlinear problems where the correlation of the objective function is observed. Redundancy checks are important because it may ease the burden of the decision maker if redundant objectives are eliminated.

3. THEORETICAL BACKGROUND

We present a set of optimality conditions for multiobjective optimization problems. Because the conditions are different for differentiable and nondifferentiable problem, they are handled separately.

3.1. Differentiable Optimality Conditions

Optimality conditions are an important sector in optimization. As elsewhere, we restrict the treatment also here to finite-dimensional Euclidean spaces. We consider problems of the form

(3.1.1)
$$\begin{aligned} &\text{minimize} && \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\} \\ &\text{subject to} && \mathbf{x} \in S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}. \end{aligned}$$

We denote the set of active constraints at a point $\mathbf{x}^*$ by

$$J(\mathbf{x}^*) = \{j \in \{1, \ldots, m\} \mid g_j(\mathbf{x}^*) = 0\}.$$

We assume in this section that the objective and the constraint functions are continuously differentiable. In Section 3.2 we treat nondifferentiable functions.

Similar optimality results are also handled, for example, in Da Cunha and Polak (1967), Kuhn and Tucker (1951), Marusciac (1982), Simon (1986) and Yu (1985, pp. 35–38, 49–50). In order to highlight the ideas, the theorems are here presented in a simplified form as compared to the general practice. For this reason, the proofs have been modified.

3.1.1. First-Order Conditions

We begin with a necessary condition of the Fritz John type.

Theorem 3.1.1. (*Fritz John necessary condition for Pareto optimality*) Let the objective and the constraint functions of problem (3.1.1) be continuously differentiable at a decision vector $\mathbf{x}^* \in S$. A necessary condition for $\mathbf{x}^*$ to be

Pareto optimal is that there exist vectors $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ for which $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$ such that

$$(1) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m.$$

Proof. See, for instance, Da Cunha and Polak (1967).

We do not present the proof here because it is quite extensive. The theorem can be considered a special case of the corresponding theorem for nondifferentiable problems, which is proved in Subsection 3.2.1. For convex problems, necessary optimality conditions can be derived by using separating hyperplanes. This is realized, for example, in Zadeh (1963). A separation theorem is also employed in the proof of the general case in Da Cunha and Polak (1967).

Corollary 3.1.2. (*Fritz John necessary condition for weak Pareto optimality*) The condition of Theorem 3.1.1 is also necessary for a decision vector $\mathbf{x}^* \in S$ to be weakly Pareto optimal.

The difference between Fritz John type and Karush-Kuhn-Tucker type optimality conditions in single objective optimization is that the multiplier (λ) of the objective function is assumed to be positive in the latter case. This eliminates degeneracy since it implies that the objective function plays its important role in the optimality conditions. To guarantee the positivity of λ, some regularity has to be assumed in the problem. Different regularity conditions exist and they are called *constraint qualifications*.

In the multiobjective case it is equally important that all the multipliers of the objective functions are not equal to zero. Sometimes the multipliers connected to Karush-Kuhn-Tucker optimality conditions are called *Karush-Kuhn-Tucker multipliers*. This concept will be used later.

In order to present the Karush-Kuhn-Tucker optimality conditions we must formulate some constraint qualification. From among several different alternatives we here present the so-called Kuhn-Tucker constraint qualification.

Definition 3.1.3. Let the constraint functions g_j of problem (3.1.1) be continuously differentiable at $\mathbf{x}^* \in S$. The problem satisfies the *Kuhn-Tucker constraint qualification* at $\mathbf{x}^*$ if for any $\mathbf{d} \in \mathbf{R}^n$ such that $\nabla g_j(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $j \in J(\mathbf{x}^*)$, there exists a function $\mathbf{a}\colon [0,1] \to \mathbf{R}^n$ which is continuously differentiable at 0, and some real scalar $\alpha > 0$, such that

$$\mathbf{a}(0) = \mathbf{x}^*, \quad \mathbf{g}(\mathbf{a}(t)) \leq \mathbf{0} \quad \text{for all} \quad 0 \leq t \leq 1 \quad \text{and} \quad \mathbf{a}'(0) = \alpha \mathbf{d}.$$

Before we can continue, we write down the so-called Motzkin's theorem of the alternative. It will be needed in the proof of the following necessary condition.

Theorem 3.1.4. (*Motzkin's theorem*) Let A and C be given matrices. Then either the system of inequalities

$$A\mathbf{x} < \mathbf{0}, \quad C\mathbf{x} \leq \mathbf{0}$$

has a solution $\mathbf{x}$, or the system

$$A^T\boldsymbol{\lambda} + C^T\boldsymbol{\mu} = 0, \quad \boldsymbol{\lambda} \geq \mathbf{0}, \quad \boldsymbol{\lambda} \neq \mathbf{0}, \quad \boldsymbol{\mu} \geq \mathbf{0}$$

has a solution $(\boldsymbol{\lambda}, \boldsymbol{\mu})$, but never both.

Proof. See, for example, Mangasarian (1969, pp. 28–29).

Now we can formulate the Karush-Kuhn-Tucker necessary condition for Pareto optimality.

Theorem 3.1.5. (*Karush-Kuhn-Tucker necessary condition for Pareto optimality*) Let the assumptions of Theorem 3.1.1 be satisfied by the Kuhn-Tucker constraint qualification. Theorem 3.1.1 is then valid with the addition that $\boldsymbol{\lambda} \neq \mathbf{0}$.

Proof. Let $\mathbf{x}^* \in S$ be Pareto optimal. The idea of this proof is to apply Theorem 3.1.4. For this reason we prove that there does not exist any $\mathbf{d} \in \mathbf{R}^n$ such that

$$\begin{aligned} &\nabla f_i(\mathbf{x}^*)^T\mathbf{d} < 0 \quad \text{for all} \quad i = 1, \ldots, k, \quad \text{and} \\ &\nabla g_j(\mathbf{x}^*)^T\mathbf{d} \leq 0 \quad \text{for all} \quad j \in J(\mathbf{x}^*). \end{aligned} \tag{3.1.2}$$

Let us on the contrary assume that there exists some $\mathbf{d}^* \in \mathbf{R}^n$ satisfying (3.1.2). Then from the Kuhn-Tucker constraint qualification we know that there exists a function $\mathbf{a}\colon [0,1] \to \mathbf{R}^n$ which is continuously differentiable at 0 and some real scalar $\alpha > 0$ such that $\mathbf{a}(0) = \mathbf{x}^*$, $\mathbf{g}(\mathbf{a}(t)) \leq \mathbf{0}$ for all $0 \leq t \leq 1$ and $\mathbf{a}'(0) = \alpha\mathbf{d}^*$.

Because the functions f_i are continuously differentiable, we can approximate $f_i(\mathbf{a}(t))$ linearly as

$$\begin{aligned} f_i(\mathbf{a}(t)) &= f_i(\mathbf{x}^*) + \nabla f_i(\mathbf{x}^*)^T(\mathbf{a}(t) - \mathbf{x}^*) + \|\mathbf{a}(t) - \mathbf{x}^*\|\varphi(\mathbf{a}(t), \mathbf{x}^*) \\ &= f_i(\mathbf{x}^*) + \nabla f_i(\mathbf{x}^*)^T(\mathbf{a}(t) - \mathbf{a}(0)) + \|\mathbf{a}(t) - \mathbf{a}(0)\|\varphi(\mathbf{a}(t), \mathbf{a}(0)) \\ &= f_i(\mathbf{x}^*) + t\nabla f_i(\mathbf{x}^*)^T\Big(\frac{\mathbf{a}(0+t) - \mathbf{a}(0)}{t}\Big) + \|\mathbf{a}(t) - \mathbf{a}(0)\|\varphi(\mathbf{a}(t), \mathbf{a}(0)), \end{aligned}$$

where $\varphi(\mathbf{a}(t), \mathbf{a}(0)) \to 0$ as $\|\mathbf{a}(t) - \mathbf{a}(0)\| \to 0$. As $t \to 0$ tends $\|\mathbf{a}(t) - \mathbf{a}(0)\|$ to zero and $(\mathbf{a}(0+t) - \mathbf{a}(0))/t \to \mathbf{a}'(0) = \alpha\mathbf{d}^*$.

After utilizing the assumption $\nabla f_i(\mathbf{x}^*)^T \mathbf{d}^* < 0$ for all $i = 1, \ldots, k$ (and $t \geq 0$), we have $f_i(\mathbf{a}(t)) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ for a sufficiently small t. This contradicts the Pareto optimality of $\mathbf{x}^*$.

Thus we have proved statement (3.1.2). Now we conclude from Theorem 3.1.4 that there exist multipliers $\lambda_i \geq 0$ for $i = 1, \ldots, k$, $\boldsymbol{\lambda} \neq \mathbf{0}$, and $\mu_j \geq 0$ for $j \in J(\mathbf{x}^*)$ such that $\sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j \in J(\mathbf{x}^*)} \mu_j \nabla g_j(\mathbf{x}^*) = 0$. We obtain statement (1) of Theorem 3.1.1 by setting $\mu_j = 0$ for all $j \in \{1, \ldots, m\} \setminus J(\mathbf{x}^*)$.

If $g_j(\mathbf{x}^*) < 0$ for some $j = 1, \ldots, m$, then according to the above setting $\mu_j = 0$ and equalities (2) of Theorem 3.1.1 follow. □

A proof basically similar but different in realization is presented in Marusciac (1982).

Corollary 3.1.6. (*Karush-Kuhn-Tucker necessary condition for weak Pareto optimality*) The condition of Theorem 3.1.5 is also necessary for a decision vector $\mathbf{x}^* \in S$ to be weakly Pareto optimal.

Constraint qualifications based on the linear independence of gradient vectors are stated in Da Cunha and Polak (1967). Other constraint qualifications are collected in Simon (1986). In addition, a new constraint qualification for convex problems is introduced in Zhou et al. (1993).

If the multiobjective optimization problem is convex, then we can state a sufficient condition for Pareto optimality. Let us first recall the sufficient condition of optimality in the single objective case.

Theorem 3.1.7. (*Karush-Kuhn-Tucker sufficient condition for optimality*) A sufficient condition for a point $\mathbf{x}^* \in \mathbf{R}^n$ to be a (global) minimum of the problem

$$\begin{aligned} &\text{minimize} && f_i(\mathbf{x}) \\ &\text{subject to} && \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}, \end{aligned}$$

where the objective function $f_i \colon \mathbf{R}^n \to \mathbf{R}$ and the constraints $g_j \colon \mathbf{R}^n \to \mathbf{R}$, $j = 1, \ldots, m$, are convex and continuously differentiable at $\mathbf{x}^*$, is that there exist multipliers $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that

$$(1) \quad \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m.$$

Proof. See, for example, Simon (1986).

Now we can extend Theorem 3.1.7 for the multiobjective case.

Theorem 3.1.8. (*Karush-Kuhn-Tucker sufficient condition for Pareto optimality*) Let the objective and the constraint functions of problem (3.1.1) be

convex and continuously differentiable at a decision vector $\mathbf{x}^* \in S$. A sufficient condition for $\mathbf{x}^*$ to be Pareto optimal is that there exist multipliers $\mathbf{0} < \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that

$$(1) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \text{ for all } j = 1, \ldots, m.$$

Proof. Let the vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ be such that the conditions stated are satisfied. We define a function $F\colon \mathbf{R}^n \to \mathbf{R}$ as $F(\mathbf{x}) = \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x})$, where $\mathbf{x} \in S$. Trivially F is convex because all the functions f_i are and we have $\boldsymbol{\lambda} > \mathbf{0}$. Now from statements (1) and (2), we obtain $\nabla F(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$ and $\mu_j g_j(\mathbf{x}^*) = 0$ for all $j = 1, \ldots, m$. Thus, according to Theorem 3.1.7, the sufficient condition for F to attain its minimum at $\mathbf{x}^*$ is satisfied. So $F(\mathbf{x}^*) \leq F(\mathbf{x})$ for all $\mathbf{x} \in S$. In other words,

$$(3.1.3) \qquad \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}^*) \leq \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x})$$

for all $\mathbf{x} \in S$.

Let us assume that $\mathbf{x}^*$ is not Pareto-optimal. Then there exists some point $\hat{\mathbf{x}} \in S$ such that $f_i(\hat{\mathbf{x}}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and for at least one index j is $f_j(\hat{\mathbf{x}}) < f_j(\mathbf{x}^*)$. Because every λ_i was assumed to be positive, we have $\sum_{i=1}^{k} \lambda_i f_i(\hat{\mathbf{x}}) < \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}^*)$. This is a contradiction with inequality (3.1.3) and $\mathbf{x}^*$ is thus Pareto optimal. □

Note that because the multiobjective optimization problem is assumed to be convex, Theorem 3.1.8 provides a sufficient condition for global Pareto optimality. This was stated in Theorem 2.2.3.

Theorem 3.1.9. (*Karush-Kuhn-Tucker sufficient condition for weak Pareto optimality*) The condition in Theorem 3.1.8 is sufficient for a decision vector $\mathbf{x}^* \in S$ to be weakly Pareto optimal for $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ with $\boldsymbol{\lambda} \neq \mathbf{0}$.

Proof. The proof is a straightforward modification of the proof of Theorem 3.1.8.

The convexity assumption in Theorem 3.1.8 can be relaxed. The stated sufficient condition is also valid if the objective functions are pseudoconvex and the constraint functions are quasiconvex. This extension is handled, for example, in Majumdar (1997), Marusciac (1982) and Simon (1986).

If an ordering cone D is used in defining efficiency, then the optimality conditions are similar to those presented above except for the multipliers λ_i. Now

they are not only nonnegative real scalars but belong to a dual cone D^*, where $D^* = \{\boldsymbol{\lambda} \in \mathbf{R}^k \mid \boldsymbol{\lambda}^T \mathbf{y} \geq 0 \text{ for all } \mathbf{y} \in D\}$. Because of the close resemblance, we do not here handle optimality conditions separately for efficiency. For details see, for example, Chen (1984) and Luc (1989, pp. 74–79).

3.1.2. Second-Order Conditions

Second-order optimality conditions (presuming twice continuously differentiable objective and constraint functions) have been examined substantially less than first-order optimality conditions. Second-order optimality conditions provide a means of reducing the set of candidate solutions produced by the first-order conditions but at the same time tighten the assumptions set to the regularity of the problem.

Second-order optimality conditions for (local) Pareto optimality are treated, for example, in Wan (1975). For completeness, we here present examples of necessary and sufficient second-order optimality conditions following Wang (1991).

First we need one more constraint qualification, namely the regularity of decision vectors.

Definition 3.1.10. A point $\mathbf{x}^* \in S$ is said to be a *regular point* if the gradients of the active constraints at $\mathbf{x}^*$ are linearly independent.

Theorem 3.1.11. (*Second-order necessary condition for Pareto optimality*) Let the objective and the constraint functions of problem (3.1.1) be twice continuously differentiable at a regular decision vector $\mathbf{x}^* \in S$. A necessary condition for $\mathbf{x}^*$ to be Pareto optimal is that there exist vectors $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$, $\boldsymbol{\lambda} \neq \mathbf{0}$, and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that

$$(1) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m$$

$$(3) \quad \mathbf{d}^T \left(\sum_{i=1}^{k} \lambda_i \nabla^2 f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla^2 g_j(\mathbf{x}^*) \right) \mathbf{d} \geq 0$$

for all $\mathbf{d} \in \{\mathbf{0} \neq \mathbf{d} \in \mathbf{R}^n \mid \nabla f_i(\mathbf{x}^*)^T \mathbf{d} \leq 0 \text{ for all } i = 1, \ldots, k, \ \nabla g_j(\mathbf{x}^*)^T \mathbf{d} = 0 \text{ for all } j \in J(\mathbf{x}^*)\}$.

Proof. See Wang (1991).

Note that when second-order optimality conditions are concerned, we need some kind of *second-order constraint qualifications* even if we do not obtain a result satisfying $\boldsymbol{\lambda} \neq \mathbf{0}$. In Theorem 3.1.11, the regularity, that is, the linear independence of the gradients of the active constraints at the point considered,

is the kind of a second-order constraint qualification that can guarantee an even stronger result where not all the λ-coefficients can vanish.

The following theorem gives two second-order sufficient optimality conditions. The difference lies in the sets of search directions.

Theorem 3.1.12. (*Second-order sufficient condition for Pareto optimality*) Let the objective and the constraint functions of problem (3.1.1) be twice continuously differentiable at a decision vector $\mathbf{x}^* \in S$. A sufficient condition for $\mathbf{x}^*$ to be Pareto optimal is that there exist vectors $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ for which $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$ such that

$$(1) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m$$

$$(3) \quad \mathbf{d}^T \Big(\sum_{i=1}^{k} \lambda_i \nabla^2 f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla^2 g_j(\mathbf{x}^*) \Big) \mathbf{d} > 0$$

for either all $\mathbf{d} \in \{\mathbf{0} \neq \mathbf{d} \in \mathbf{R}^n \mid \nabla f_i(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $i = 1, \ldots, k$, $\nabla g_j(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $j \in J(\mathbf{x}^*)\}$ or all $\mathbf{d} \in \{\mathbf{0} \neq \mathbf{d} \in \mathbf{R}^n \mid \nabla g_j(\mathbf{x}^*)^T \mathbf{d} = 0$ for all $j \in J^+(\mathbf{x}^*)$, $\nabla g_j(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $j \in J(\mathbf{x}^*) \setminus J^+(\mathbf{x}^*)\}$, where $J^+(\mathbf{x}^*) = \{j \in J(\mathbf{x}^*) \mid \mu_j > 0\}$.

Proof. See Wang (1991).

Second-order sufficient conditions for Pareto optimality are also treated in Simon (1986), and more necessary and sufficient conditions for Pareto and weakly Pareto optimal solutions are presented in Wang (1991).

3.1.3. Conditions for Proper Pareto Optimality

For completeness we also present the original necessary optimality condition formulated for proper Pareto optimality in the sense of Kuhn and Tucker (see Definition 2.9.3) as stated by Kuhn and Tucker (1951). To begin with, we write down Tucker's theorem of the alternative, which will be utilized in the proof.

Theorem 3.1.13. (*Tucker's theorem*) Let A and C be given matrices. Then either the system of inequalities

$$A\mathbf{x} \leq \mathbf{0}, \quad A\mathbf{x} \neq \mathbf{0}, \quad C\mathbf{x} \leq \mathbf{0}$$

has a solution $\mathbf{x}$, or the system

$$A^T\boldsymbol{\lambda} + C^T\boldsymbol{\mu} = 0, \quad \boldsymbol{\lambda} > \mathbf{0}, \quad \boldsymbol{\mu} \geq \mathbf{0}$$

has a solution $(\boldsymbol{\lambda}, \boldsymbol{\mu})$, but never both.

Proof. The proof is similar to the proof of Theorem 3.1.4.

We can now present the necessary condition for proper Pareto optimality.

Theorem 3.1.14. (*Kuhn-Tucker necessary condition for proper Pareto optimality*) Let the objective and the constraint functions of problem (3.1.1) be continuously differentiable at a decision vector $\mathbf{x}^* \in S$. A necessary condition for $\mathbf{x}^*$ to be properly Pareto optimal (in the sense of Kuhn and Tucker) is that there exist vectors $\mathbf{0} < \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that

$$(1) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m.$$

Proof. Let $\mathbf{x}^*$ be properly Pareto optimal (in the sense of Kuhn and Tucker). From the definition we know that no vector $\mathbf{d} \in \mathbf{R}^n$ exists such that $\nabla f_i(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $i = 1, \ldots, k$, $\nabla f_j(\mathbf{x}^*)^T \mathbf{d} < 0$ for some index j, and $\nabla g_l(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $l \in J(\mathbf{x}^*)$. Then, from Theorem 3.1.13 we know that there exist multipliers $\lambda_i > 0$ for $i = 1, \ldots, k$ and $\mu_j \geq 0$ for $j \in J(\mathbf{x}^*)$ such that $\sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j \in J(\mathbf{x}^*)} \mu_j \nabla g_j(\mathbf{x}^*) = 0$. We obtain statement (1) by setting $\mu_j = 0$ for all $j \in \{1, \ldots, m\} \setminus J(\mathbf{x}^*)$.

If $g_j(\mathbf{x}^*) < 0$ for some j, then according to the above setting $\mu_j = 0$ and equalities (2) follow. □

It is proved in Geoffrion (1968) and Sawaragi et al. (1985, p. 90), that if the Kuhn-Tucker constraint qualification (Definition 3.1.3) is satisfied at a decision vector $\mathbf{x}^* \in S$, then the condition in Theorem 3.1.14 is also necessary for $\mathbf{x}^*$ to be properly Pareto optimal in the sense of Geoffrion. Finally, we write down the sufficient condition for proper Pareto optimality.

Theorem 3.1.15. (*Kuhn-Tucker sufficient condition for proper Pareto optimality*) If problem (3.1.1) is convex, then the condition in Theorem 3.1.14 is also sufficient for a decision vector $\mathbf{x}^* \in S$ to be properly Pareto optimal (in the sense of Kuhn and Tucker).

Proof. See Sawaragi et al. (1985, p. 90) or Shimizu et al. (1997, p. 112).

Let us finally mention that necessary and sufficient conditions for proper Pareto optimality in the sense of Geoffrion are presented in Gulati and Islam (1990) for pseudolinear objective (i.e., differentiable functions that are both pseudoconvex and pseudoconcave) and quasiconvex constraint functions.

3.2. Nondifferentiable Optimality Conditions

In this section, we no longer necessitate differentiability but put forward nondifferentiable counterparts for the optimality conditions presented in Section 3.1. Usually, when the assumption of continuous differentiability is given up, functions are assumed to be locally Lipschitzian (see Definition 2.1.12). Remember that a function is here called nondifferentiable if it is locally Lipschitzian (and not necessarily continuously differentiable).

In every other way the multiobjective optimization problem to be solved is still of the form

(3.2.1)
$$\begin{aligned} &\text{minimize} && \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\} \\ &\text{subject to} && \mathbf{x} \in S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}. \end{aligned}$$

We first briefly present some properties of subdifferentials (see Definition 2.1.14) without any proofs.

Theorem 3.2.1. Let the functions $f_i\colon \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, k$, be locally Lipschitzian at a point $\mathbf{x}^* \in \mathbf{R}^n$. Then, for weights $w_i \in \mathbf{R}$ we have

$$\partial\Big(\sum_{i=1}^{k} w_i f_i\Big)(\mathbf{x}^*) \subset \sum_{i=1}^{k} w_i \partial f_i(\mathbf{x}^*).$$

The two sets are equal if at least $k - 1$ of the functions f_i are continuously differentiable, or if the functions are convex and the weights are positive.

Proof. See, for example, Mäkelä and Neittaanmäki (1992, p. 39) and Clarke (1983, pp. 38–39).

Theorem 3.2.2. Let the functions $f_i\colon \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, k$, be locally Lipschitzian at a point $\mathbf{x}^* \in \mathbf{R}^n$. Then the function $f\colon \mathbf{R}^n \to \mathbf{R}$

$$f(\mathbf{x}) = \max_{i=1,\ldots,k} f_i(\mathbf{x})$$

is also locally Lipschitzian at $\mathbf{x}^*$. In addition,

$$\partial f(\mathbf{x}^*) \subset \operatorname{conv}\{\partial f_i(\mathbf{x}^*) \mid i \in I(\mathbf{x}^*)\},$$

where $I(\mathbf{x}^*) \subset \{1, \ldots, k\}$ denotes the set of indices i for which $f(\mathbf{x}^*) = f_i(\mathbf{x}^*)$.

Proof. See, for example, Mäkelä and Neittaanmäki (1992, pp. 47–49).

Theorem 3.2.3. Let the function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ be locally Lipschitzian at a point $\mathbf{x}^* \in \mathbf{R}^n$ and attain its (local) minimum at $\mathbf{x}^*$, then

$$0 \in \partial f_i(\mathbf{x}^*).$$

If the function f_i is convex, then the condition is also sufficient and the minimum is global.

Proof. See, for example, Mäkelä and Neittaanmäki (1992, pp. 70–71).

Before moving on to the optimality conditions of the Fritz John and Karush-Kuhn-Tucker type we should point out the following. If a single objective function is defined on a set, the counterpart of the condition in Theorem 3.2.3 says that zero belongs to the algebraic sum of two sets formed at the point considered. The sets are the subdifferential of the objective function and the normal cone of the feasible region. This result is adapted for convex multiobjective optimization problems involving continuous objective functions and closed feasible regions in Plastria and Carrizosa (1996). The necessary and sufficient condition for weak Pareto optimality is that zero belongs to the sum of the union of the subdifferentials of the objective functions and the normal cone of the feasible region. Note that the functions do not have to be even locally Lipschitzian. According to Clarke (1983, pp. 230–231), the same condition is necessary for weak Pareto optimality in general problems as well. We do not treat these results more thoroughly here. Instead, we present one more result for single objective nondifferentiable optimization.

Theorem 3.2.4. (*Fritz John necessary condition for optimality*) A necessary condition for a point $\mathbf{x}^* \in \mathbf{R}^n$ to be a local minimum of the problem

$$\begin{aligned} &\text{minimize} && f_i(\mathbf{x}) \\ &\text{subject to} && \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}, \end{aligned}$$

where the objective function $f_i\colon \mathbf{R}^n \to \mathbf{R}$ and the constraints $g_j\colon \mathbf{R}^n \to \mathbf{R}$, $j = 1, \ldots, m$, are locally Lipschitzian at $\mathbf{x}^*$, is that there exist multipliers $0 \leq \lambda \in \mathbf{R}$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ for which $(\lambda, \boldsymbol{\mu}) \neq (0, \mathbf{0})$ such that

$$(1) \quad 0 \in \lambda \partial f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \partial g_j(\mathbf{x}^*)$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m.$$

Proof. See, for example, Clarke (1983, pp. 228–230) or Kiwiel (1985c, p. 16).

Now that we have assembled a set of tools, we are in a position to handle the actual optimality conditions. More information can be found, for example, in Craven (1989), Doležal (1985), Minami (1980–81, 1981, 1983), Shimizu et al. (1997, pp. 322–325) and Wang (1984). The theorems are presented in a simplified form here compared to the general practice so as to emphasize the ideas. For this reason, the proofs have been modified.

3.2.1. First-Order Conditions

The first result to be presented is a necessary condition of the Fritz John type for Pareto optimality.

Theorem 3.2.5. (*Fritz John necessary condition for Pareto optimality*) Let the objective and the constraint functions of problem (3.2.1) be locally Lipschitzian at a point $\mathbf{x}^* \in S$. A necessary condition for the point $\mathbf{x}^*$ to be Pareto optimal is that there exist multipliers $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ for which $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$ such that

$$(1) \quad 0 \in \sum_{i=1}^{k} \lambda_i \partial f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \partial g_j(\mathbf{x}^*)$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m.$$

Proof. Because it is assumed that $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$, we can normalize the multipliers to sum up to one. We shall here prove a stronger condition, where $\sum_{i=1}^{k} \lambda_i + \sum_{j=1}^{m} \mu_j = 1$.

Let $\mathbf{x}^* \in S$ be Pareto optimal. At first we define an additional function $F\colon \mathbf{R}^n \to \mathbf{R}$ by

$$F(\mathbf{x}) = \max\left[f_i(\mathbf{x}) - f_i(\mathbf{x}^*), g_j(\mathbf{x}) \mid i = 1, \ldots, k, \ j = 1, \ldots, m\right]$$

and show that for all $\mathbf{x} \in \mathbf{R}^n$ is

$$F(\mathbf{x}) \geq 0. \tag{3.2.2}$$

Let us on the contrary assume that for some $\mathbf{x}^\circ \in \mathbf{R}^n$ is $F(\mathbf{x}^\circ) < 0$. Then $g_j(\mathbf{x}^\circ) < 0$ for all $j = 1, \ldots, m$ and the point $\mathbf{x}^\circ$ is thus feasible in problem (3.2.1). In addition, $f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$, which contradicts the Pareto optimality of $\mathbf{x}^*$. Thus (3.2.2) must be true.

Noting that the point $\mathbf{x}^*$ is feasible in problem (3.2.1), we obtain $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$. This implies $F(\mathbf{x}^*) = 0$. Combining this fact with property (3.2.2), we know that F attains its (global) minimum at $\mathbf{x}^*$. As all the functions f_i and g_j are locally Lipschitzian at $\mathbf{x}^*$, likewise F (according to Theorem 3.2.2). We deduce from Theorem 3.2.3 that $0 \in \partial F(\mathbf{x}^*)$.

Note that

$$\partial(f_i(\mathbf{x}) - f_i(\mathbf{x}^*)) = \partial f_i(\mathbf{x}), \tag{3.2.3}$$

applying Theorem 3.2.1.

We designate the set of indices j for which $F(\mathbf{x}^*) = g_j(\mathbf{x}^*)$ by $J(\mathbf{x}^*) \subset \{1, \ldots, m\}$. Now we can employ Theorem 3.2.2 and (3.2.3) and obtain

$$0 \in \operatorname{conv}\left\{\partial f_i(\mathbf{x}^*), \partial g_j(\mathbf{x}^*) \mid i = 1, \ldots, k, \ j \in J(\mathbf{x}^*)\right\}.$$

Employing the definition of a convex hull, we know that there exist vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ of real multipliers for which $\lambda_i \geq 0$ for all $i = 1, \ldots, k$, $\mu_j \geq 0$ for all $j \in J(\mathbf{x}^*)$ and $\sum_{i=1}^{k} \lambda_i + \sum_{j \in J(\mathbf{x}^*)} \mu_j = 1$, such that

$$0 \in \sum_{i=1}^{k} \lambda_i \partial f_i(\mathbf{x}^*) + \sum_{j \in J(\mathbf{x}^*)} \mu_j \partial g_j(\mathbf{x}^*).$$

Now we can set $\mu_j = 0$ for all $j \in \{1, \ldots, m\} \setminus J(\mathbf{x}^*)$. Statement (1) follows from this setting.

Part (2) is trivial. If $g_j(\mathbf{x}^*) < 0$ for some j, then $j \in \{1, \ldots, m\} \setminus J(\mathbf{x}^*)$ and we have $\mu_j = 0$. This completes the proof. □

Now we can define decision vectors called substationary points.

Definition 3.2.6. A decision vector $\mathbf{x}^* \in S$ is called a *substationary point* if it satisfies the (necessary) optimality condition presented in Theorem 3.2.5.

Theorem 3.2.5 can also be proved by first employing a scalarization method and then Theorem 3.2.4 for the resulting single objective optimization problem (see, e.g., Doležal (1985)).

Corollary 3.2.7. (*Fritz John necessary condition for weak Pareto optimality*) The condition of Theorem 3.2.5 is also necessary for a decision vector $\mathbf{x}^* \in S$ to be weakly Pareto optimal.

Next, we examine some constraint qualifications. It is obvious that they differ from the differentiable case.

Note that when the necessary optimality conditions are derived with the help of a scalarizing function, it is easy to generalize the constraint qualifications from single objective optimization to the multiobjective case. One simply assumes that both the original constraints and the possible additional constraints satisfy a constraint qualification. This is expressed in Doležal (1985). The constraint qualifications used there are those of calmness and Mangasarian-Fromovitz.

The so-called Cottle constraint qualification is used in the following theorem. Other constraint qualifications are presented, for example, in Ishizuka and Shimizu (1984).

Definition 3.2.8. Let the objective and the constraint functions of problem (3.2.1) be locally Lipschitzian at a point $\mathbf{x}^* \in S$. Problem (3.2.1) satisfies the *Cottle constraint qualification* at $\mathbf{x}^*$ if either $g_j(\mathbf{x}^*) < 0$ for all $j = 1, \ldots, m$ or $0 \notin \operatorname{conv} \{\partial g_j(\mathbf{x}^*) \mid g_j(\mathbf{x}^*) = 0\}$.

Assuming the Cottle constraint qualification, we obtain the Karush-Kuhn-Tucker necessary condition for Pareto optimality.

Theorem 3.2.9. (*Karush-Kuhn-Tucker necessary condition for Pareto optimality*) Let the assumptions of Theorem 3.2.5 be satisfied by the Cottle constraint qualification. Theorem 3.2.5 is then valid with the addition that $\boldsymbol{\lambda} \neq \mathbf{0}$.

Proof. The proof of Theorem 3.2.5 is here valid up to the observation $0 \in \partial F(\mathbf{x}^*)$ and result (3.2.3). We prove also this theorem in a stronger form, where the multipliers sum up to one.

From the definition of F we know that

$$F(\mathbf{x}^*) = 0.$$

We continue by first assuming that $g_j(\mathbf{x}^*) < 0$ for all $j = 1, \ldots, m$. In this case, $F(\mathbf{x}^*) > g_j(\mathbf{x}^*)$ for all j. Now we can apply Theorem 3.2.2 and equation (3.2.3) and obtain

$$0 \in \operatorname{conv}\{\partial f_i(\mathbf{x}^*) \mid i = 1, \ldots, k\}.$$

From the definition of a convex hull we know that there exists a vector $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ of multipliers for which $\sum_{i=1}^k \lambda_i = 1$ (thus $\boldsymbol{\lambda} \neq \mathbf{0}$) such that

$$0 \in \sum_{i=1}^k \lambda_i \partial f_i(\mathbf{x}^*).$$

We obtain the statement to be proved (denoted by (1) in Theorem 3.2.5) by setting $\mu_j = 0$ for all $j = 1, \ldots, m$.

On the other hand, if there exists some index j such that $g_j(\mathbf{x}^*) = 0$, we denote the set of such indices by $J(\mathbf{x}^*)$. By the Cottle constraint qualification we know that

$$0 \notin \operatorname{conv}\{\partial g_j(\mathbf{x}^*) \mid j \in J(\mathbf{x}^*)\}. \tag{3.2.4}$$

In this case, we deduce from Theorem 3.2.2 and result (3.2.3) that

$$0 \in \operatorname{conv}\{\partial f_i(\mathbf{x}^*), \partial g_j(\mathbf{x}^*) \mid i = 1, \ldots, k,\ j \in J(\mathbf{x}^*)\}.$$

Applying the definition of a convex hull, we know that there exist multipliers $\lambda_i \geq 0$, $i = 1, \ldots, k$, and $\mu_j \geq 0$, $j \in J(\mathbf{x}^*)$, for which $\sum_{i=1}^k \lambda_i + \sum_{j \in J(\mathbf{x}^*)} \mu_j = 1$, and by assumption (3.2.4), especially $\boldsymbol{\lambda} \neq \mathbf{0}$, such that

$$0 \in \sum_{i=1}^k \lambda_i \partial f_i(\mathbf{x}^*) + \sum_{j \in J(\mathbf{x}^*)} \mu_j \partial g_j(\mathbf{x}^*).$$

Again, we obtain the statement to be proved by setting $\mu_j = 0$ for all $j \in \{1, \ldots, m\} \setminus J(\mathbf{x}^*)$.

The proof of part (2) is the same as in Theorem 3.2.5. □

Corollary 3.2.10. (*Karush-Kuhn-Tucker necessary condition for weak Pareto optimality*) The condition of Theorem 3.2.9 is also necessary for a decision vector $\mathbf{x}^* \in S$ to be weakly Pareto optimal.

If we assume the multiobjective optimization problem to be convex and $\boldsymbol{\lambda} > \mathbf{0}$, we get a Karush-Kuhn-Tucker sufficient condition for Pareto optimality. Note that the convexity of a function implies that the function is locally Lipschitzian at any point $\mathbf{x} \in S$.

Theorem 3.2.11. (*Karush-Kuhn-Tucker sufficient condition for Pareto optimality*) Let problem (3.2.1) be convex. A sufficient condition for a decision vector $\mathbf{x}^* \in S$ to be Pareto optimal is that there exist multipliers $\mathbf{0} < \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that

$$(1) \quad 0 \in \sum_{i=1}^{k} \lambda_i \partial f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \partial g_j(\mathbf{x}^*)$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m.$$

Proof. To start with, we define an additional function $F\colon \mathbf{R}^n \to \mathbf{R}$ by $F(\mathbf{x}) = \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}) + \sum_{j=1}^{m} \mu_j g_j(\mathbf{x})$, where the multipliers λ_i and μ_j satisfy the above assumptions. Because the functions f_i and g_j are convex, $\boldsymbol{\lambda} > \mathbf{0}$ and $\boldsymbol{\mu} \geq \mathbf{0}$, then F too is convex, and $\partial F(\mathbf{x}) = \sum_{i=1}^{k} \lambda_i \partial f_i(\mathbf{x}) + \sum_{j=1}^{m} \mu_j \partial g_j(\mathbf{x})$ (as stated in Theorem 3.2.1).

From assumption (1) we know that $0 \in \partial F(\mathbf{x}^*)$, and, according to Theorem 3.2.3, the point $\mathbf{x}^*$ is a (global) minimum of F. This implies that for any $\mathbf{x}^\circ \in \mathbf{R}^n$, especially any $\mathbf{x}^\circ$ satisfying $\mathbf{g}(\mathbf{x}^\circ) \leq \mathbf{0}$, the following is valid:

$$\begin{aligned} 0 &\leq F(\mathbf{x}^\circ) - F(\mathbf{x}^*) \\ &= \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}^\circ) + \sum_{j=1}^{m} \mu_j g_j(\mathbf{x}^\circ) - \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}^*) - \sum_{j=1}^{m} \mu_j g_j(\mathbf{x}^*). \end{aligned}$$

Employing assumption (2), the fact that $\mathbf{g}(\mathbf{x}^\circ) \leq \mathbf{0}$ and $\boldsymbol{\mu} \geq \mathbf{0}$, we obtain

$$\sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}^*) \leq \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}^\circ) \tag{3.2.5}$$

for any $\mathbf{x}^\circ \in S$.

Let us assume that $\mathbf{x}^*$ is not Pareto-optimal. Then there exists some feasible $\hat{\mathbf{x}}$ such that $f_i(\hat{\mathbf{x}}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and for at least one index j is $f_j(\hat{\mathbf{x}}) < f_j(\mathbf{x}^*)$. Because every λ_i was assumed to be positive, we have $\sum_{i=1}^{k} \lambda_i f_i(\hat{\mathbf{x}}) < \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}^*)$. This contradicts inequality (3.2.5) and $\mathbf{x}^*$ is thus Pareto optimal. □

Theorem 3.2.12. (*Karush-Kuhn-Tucker sufficient condition for weak Pareto optimality*) The condition stated in Theorem 3.2.11 is sufficient for a decision vector $\mathbf{x}^* \in S$ to be weakly Pareto optimal for $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ with $\boldsymbol{\lambda} \neq \mathbf{0}$.

Proof. The proof is a trivial modification of the proof of Theorem 3.2.11.

Finally, we introduce one more constraint qualification. It can only be applied to convex problems and it will also be needed in Part II (Section 2.2) in connection with the multiobjective proximal bundle method. It is called the Slater constraint qualification. It is independent of the differentiability of the functions involved.

Definition 3.2.13. Let problem (3.2.1) be convex. Problem (3.2.1) satisfies the *Slater constraint qualification* if there exists some $\mathbf{x}$ with $g_j(\mathbf{x}) < 0$ for all $j = 1, \ldots, m$.

Theorem 3.2.9 and Corollary 3.2.10 can now be reformulated for convex problems assuming the Slater constraint qualification. Remember that convexity implies that functions are locally Lipschitzian at any point in the feasible region.

Theorem 3.2.14. (*Karush-Kuhn-Tucker necessary condition for (weak) Pareto optimality*) Let problem (3.2.1) be convex, satisfying the Slater constraint qualification. A necessary condition for a point $\mathbf{x}^* \in S$ to be (weakly) Pareto optimal is that there exist multipliers $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ with $\boldsymbol{\lambda} \neq \mathbf{0}$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that

$$(1) \quad 0 \in \sum_{i=1}^{k} \lambda_i \partial f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \partial g_j(\mathbf{x}^*)$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m.$$

Proof. The proof is a trivial modification of the proof of Theorem 3.2.9 when we note the following. In case the set $J(\mathbf{x}^*)$ is nonempty we denote $g(\mathbf{x}) = \max[g_j(\mathbf{x}) \mid j = 1, \ldots, m]$. Now $g(\mathbf{x}^*) = g_j(\mathbf{x}^*)$ for $j \in J(\mathbf{x}^*)$. By the Slater constraint qualification there exists some $\mathbf{x}^\circ$ such that $g_j(\mathbf{x}^\circ) < 0$ for all j. Thus, $\mathbf{x}^*$ cannot be the global minimum of the convex function g. According to Theorem 3.2.3 we derive

$$0 \notin \operatorname{conv}\{\partial g_j(\mathbf{x}^*) \mid j \in J(\mathbf{x}^*)\}.$$

The proof of Theorem 3.2.9 can now be applied. □

Necessary optimality conditions for Pareto optimality in those nondifferentiable problems where the objective functions are fractions of convex and

concave functions are formulated in Bhatia and Datta (1985). In addition, necessary Fritz John and Karush-Kuhn-Tucker type optimality conditions for weak Pareto optimality involving so-called semidifferentiable pre-invex functions are treated in Preda (1996).

If an ordering cone D is used in defining efficiency, then the optimality conditions are similar to those presented above, except for the multipliers λ_i (simply as in the differentiable case). The multipliers belong to the dual cone $D^* = \{\boldsymbol{\lambda} \in \mathbf{R}^k \mid \boldsymbol{\lambda}^T\mathbf{y} \geq 0 \text{ for all } \mathbf{y} \in D\}$. Because of the similarity, we do not present here separate optimality conditions for efficiency. Necessary and sufficient conditions for efficiency and weak efficiency are handled, for example, in Wang (1984). Furthermore, in Craven (1989) and El Abdouni and Thibault (1992), necessary conditions for weak efficiency in normed spaces and Banach spaces, respectively, are presented. The objective and the constraint functions are still assumed to be locally Lipschitzian.

Direct counterparts of optimality conditions for proper Pareto optimality in the sense of Kuhn and Tucker, presented in Section 3.1, cannot be stated in the nondifferentiable case. The reason is that the definition of Kuhn and Tucker assumes continuous differentiability. However, a sufficient condition for proper Pareto optimality in the sense of Geoffrion, when the objective and the constraint functions are compositions of convex, locally Lipschitzian functions, is formulated in Jeyakumar and Yang (1993). This treatment naturally includes ordinary convex, locally Lipschitzian functions. The authors also present necessary conditions for weak Pareto optimality and sufficient conditions of their own for Pareto optimality in problems with convex composite functions. A necessary and sufficient condition for proper efficiency (in the sense of Henig) is derived in Henig and Buchanan (1994, 1997) for convex problems.

3.2.2. Second-Order Conditions

At the end of this section we shall say a few words about the case where the functions involved are continuously differentiable and their gradients are locally Lipschitzian. Such functions are called $C^{1,1}$-*functions*. Second-order optimality conditions for multiobjective problems with $C^{1,1}$-functions are handled in Liu (1991). Here we briefly state the main results. First we must introduce one concept according to Liu.

Definition 3.2.15. Let the function $f_i\colon S \to \mathbf{R}$ be a $C^{1,1}$-function at the point $\mathbf{x}^* \in S$. The set

$$\partial_*^2 f_i(\mathbf{x}^*)(\mathbf{d},\mathbf{d}) = \Big\{\phi_i \in \mathbf{R} \mid \text{there exists a sequence } \{t_j\}_{j=1}^{\infty},$$
$$\phi_i = \lim_{t_j \to 0+} \frac{2}{t_j^2}\Big(f_i(\mathbf{x}^* + t_j\mathbf{d}) - f_i(\mathbf{x}^*) - t_j \nabla f_i(\mathbf{x}^*)^T\mathbf{d}\Big)\Big\}$$

is called a *generalized second-order directional derivative* of the function f_i evaluated at $\mathbf{x}^*$ in the direction $\mathbf{d} \in \mathbf{R}^n$.

The set $\partial_*^2 f_i(\mathbf{x}^*)(\mathbf{d}, \mathbf{d})$ is nonempty according to Liu and Křížek (1997).

The following second-order necessary and sufficient conditions are presented in Liu (1991), for both the case where all the objective functions are $C^{1,1}$-functions and all the constraint functions are twice continuously differentiable and the case where all the objective functions are twice continuously differentiable and all the constraint functions are $C^{1,1}$-functions. Here, we formulate only the first case, while the only difference in the conditions is the reversed roles of the Hessian matrices and generalized second-order directional derivatives.

Let us again denote the set of active constraints at $\mathbf{x}^* \in S$ by $J(\mathbf{x}^*)$.

Theorem 3.2.16. (*Second-order necessary condition for Pareto optimality*) Let the objective functions of problem (3.2.1) be $C^{1,1}$-functions and its constraint functions twice continuously differentiable at a decision vector $\mathbf{x}^* \in S$. A necessary condition for $\mathbf{x}^*$ to be Pareto optimal is that there exist vectors $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ for which $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$ such that

$$(1) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \quad \text{for all} \quad j = 1, \ldots, m$$

$$(3) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*)^T \mathbf{d} = 0, \quad \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*)^T \mathbf{d} = 0$$

$$(4) \quad \sum_{i=1}^{k} \lambda_i \phi_i + \mathbf{d}^T \left(\sum_{j=1}^{m} \mu_j \nabla^2 g_j(\mathbf{x}^*) \right) \mathbf{d} \geq 0$$

for all $\mathbf{d} \in \{\mathbf{0} \neq \mathbf{d} \in \mathbf{R}^n \mid \nabla f_i(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $i = 1, \ldots, k, \ \nabla g_j(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $j \in J(\mathbf{x}^*)\}$ and $\phi_i \in \partial_*^2 f_i(\mathbf{x}^*)(\mathbf{d}, \mathbf{d})$.

Proof. See Liu (1991).

Finally, we present a second-order sufficient optimality condition for problems involving $C^{1,1}$-functions.

Theorem 3.2.17. (*Second-order sufficient condition for Pareto optimality*) Let the objective functions of problem (3.2.1) be $C^{1,1}$-functions and its constraint functions twice continuously differentiable at a decision vector $\mathbf{x}^* \in S$. A sufficient condition for $\mathbf{x}^*$ to be Pareto optimal is that there exist vectors $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^k$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ for which $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$ such that

$$(1) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

$$(2) \quad \mu_j g_j(\mathbf{x}^*) = 0 \text{ for all } j = 1, \ldots, m$$

$$(3) \quad \sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*)^T \mathbf{d} = 0, \quad \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}^*)^T \mathbf{d} = 0$$

$$(4) \quad \sum_{i=1}^{k} \lambda_i \phi_i + \mathbf{d}^T \Big(\sum_{j=1}^{m} \mu_j \nabla^2 g_j(\mathbf{x}^*) \Big) \mathbf{d} > 0$$

for all $\mathbf{d} \in \{\mathbf{0} \neq \mathbf{d} \in \mathbf{R}^n \mid \nabla f_i(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $i = 1, \ldots, k$, $\nabla g_j(\mathbf{x}^*)^T \mathbf{d} \leq 0$ for all $j \in J(\mathbf{x}^*)\}$ and $\phi_i \in \partial_*^2 f_i(\mathbf{x}^*)(\mathbf{d}, \mathbf{d})$.

Proof. See Liu (1991).

Actually, the results in Liu (1991) are given in a more general form for efficient solutions and for problems where the constraint functions belong to a polyhedral convex cone.

3.3. More Optimality Conditions

Many necessary and sufficient conditions for weak, proper or Pareto optimality (or efficiency) have been suggested in the literature. They are based on different kinds of assumptions as to the properties and form of the problem. Many of them are based on a scalarization of the original problem and conditions are set to both the original functions and the scalarization parameters (some such conditions are presented in Part II in connection with the scalarization methods). In this book, we settle for a closer handling of the Fritz John and the Karush-Kuhn-Tucker type conditions, presented in the two earlier sections. For the interested reader we list some other references.

Necessary conditions for proper and improper Pareto optimality in the sense of Kuhn and Tucker are derived with the help of cones in Tamura and Arai (1982). Geoffrion (1968) was the first to give the basic characterization of properly Pareto optimal solutions in terms of a scalar problem, called a weighting problem (see Section 3.1 of Part II). He extended the results by a comprehensive theorem into necessary and sufficient conditions for local and global proper Pareto optimality. Geoffrion's treatment is closely followed in Chou et al. (1985), where properly Pareto optimal solutions are characterized for multiobjective optimization problems with set-valued functions. In addition, necessary and sufficient Karush-Kuhn-Tucker type optimality conditions for ϵ-Pareto optimality in convex problems using the weighting method for the objectives and exact penalty functions for the constraints are handled in Liu (1996).

In Chankong and Haimes (1982), the Karush-Kuhn-Tucker optimality conditions for Pareto optimality are modified for use in connection with certain solution methods (the ε-constraint method and the jth Lagrangian problem; see Section 3.2 of Part II). Chankong and Haimes also propose optimality conditions for proper Pareto optimality (in the sense of Geoffrion) with the ε-constraint method. Further, in Benson and Morin (1977), necessary and sufficient conditions are given for a Pareto optimal solution to be properly Pareto optimal. This is done with the help of the jth Lagrangian problem. Necessary and sufficient conditions for Pareto optimality with convex and differentiable functions partly based on the ε-constraint problem are proved in Zlobec (1984).

Necessary and sufficient conditions for Pareto optimality and proper Pareto optimality are proved with the help of duality theory and auxiliary problem (2.10.1) (presented in Section 2.10) in Wendell and Lee (1977). However, it is stated that nonlinear problems do not generally satisfy the conditions developed. In such cases Pareto optimal solutions have to be tested for proper Pareto optimality on a point-by-point basis.

In Gulati and Islam (1988), linear fractional objective functions and generalized convex constraints are handled. Necessary conditions of the Karush-Kuhn-Tucker type are presented for Pareto optimal solutions, and the conditions under which Pareto optimal solutions are properly Pareto optimal are stated. Necessary and sufficient conditions for Pareto optimality in problems with nonlinear fractional objective functions and nonlinear constraints are proved in Lee (1992). In addition, necessary optimality conditions for fractional multiobjective optimization problems with square root terms are given in Egudo (1991). In Benson (1979b), a necessary and sufficient condition is given for a point to be Pareto optimal when there are two concave objective functions (problem of maximization) and a convex feasible set.

The following references deal with conditions for efficiency, where the objective space is ordered by an ordering cone.

In Zubiri (1988), necessary and sufficient conditions are proved for weak efficiency in Banach spaces with the help of a weighted L_∞-metric (see Section 3.4 of Part II). Several necessary and sufficient conditions for efficient, weakly efficient and properly efficient solutions (in the sense of Borwein) in real topological linear spaces are collected in Jahn (1985). Necessary and sufficient optimality conditions of the Karush-Kuhn-Tucker type are derived in Hazen (1988), for cases where preferences are and are not representable by cones.

Let us finally briefly mention some further references handling nondifferentiable cases. Necessary and sufficient conditions for Pareto optimality and proper Pareto optimality are derived in Bhatia and Aggarwal (1992), by the weighting method (see Section 3.1 of Part II) and Dini derivatives. The functions in the problem are assumed to be nondifferentiable such that the objective functions are pseudoconcave and the constraint functions are quasiconvex. Some duality results are provided as well.

Optimality conditions based on the optimization theory of Dubovitskii and Milyutin presuming certain convexity assumptions are presented in Censor (1977) for Pareto optimality in $\mathbf{R}^n$ and in Minami (1981) for weak Pareto optimality in a linear topological space. No differentiability assumptions are needed. Necessary and sufficient conditions for weak, proper and Pareto optimality in finite-dimensional normed spaces are presented in Staib (1991) under different assumptions. In Shimizu et al. (1997, pp. 319–322), nondifferentiable optimality conditions assuming constraint qualifications based on directional derivatives are derived.

3.4. Sensitivity Analysis and Duality

The last topics to be mentioned in this chapter are sensitivity analysis, stability and duality. *Sensitivity analysis* studies situations when the input parameters defining the multiobjective optimization problem change or contain errors. In sensitivity analysis, an answer is sought to the question of how much the parameters can be altered and varied without affecting the solution. More justification for sensitivity analysis is provided in Rarig and Haimes (1983).

Given a family of parametrized multiobjective optimization problems, a set-valued perturbation function is defined in Tanino (1990), such that it associates with each parameter value the set of Pareto optimal points of the perturbed feasible region. The behaviour of the perturbation function is analyzed both qualitatively and quantitatively. In this context *stability* means the study of various continuity properties of the perturbation function of a family of parametrized optimization problems, that is, qualitative analysis. *Sensitivity* means the study of the derivatives of the perturbation function, that is, quantitative analysis.

In general multiobjective optimization problems, considerable attention has been paid to the stability of the preference structure of the decision maker. In these cases, it is usually assumed that the partial ordering of the objective space is induced by an ordering cone.

However, mathematical stability and sensitivity analysis are broad areas of research, and we do not intend to go into details here. Instead, we refer, for example, to Balbás and Guerra (1996), Craven (1988), Ester (1984), Gal and Wolf (1986), Kuk et al. (1996), Luc (1989), Lucchetti (1985), Papageorgiou (1985), Tanino (1988a, b, 1990) and Tanino and Sawaragi (1980), for further analysis.

Let us still mention that stability is not an unambiguous notion. As stressed, for example, in Dauer and Liu (1997), the terms and results connected to stability and sensitivity analysis are not universally defined in the literature. Different types of stability can be defined and measured in many ways. Often stability is associated with worst case performance and analysing how fast a solution degrades to a certain still acceptable level. Thus, analysis of stability is important

in implementing solutions in practice. Regardless if its significance, stability has been widely ignored in the multiple criteria decision-making context thus far.

A review of sensitivity analysis results for both linear and nonlinear problems is given in Dauer and Liu (1997). In addition, they study sensitivity analysis for MOLP problems in the objective space and deal with priority structures in goal programming. Sensitivity analysis for MOLP problems is also treated in Gal (1995).

Changes that occur in the solution of an MOLP problem, if the number of objective functions, the number of variables or the number of constraint functions is altered, are examined in Eiselt et al. (1987). This is also an interesting topic for nonlinear problems, as, for example, an objective function may have been left out of the model, and it would be useful to know how this can affect the solution obtained. For example, if a convex objective function is added to a convex multiobjective optimization problem, all weakly Pareto optimal solutions remain weakly Pareto optimal (see Lowe et al. (1984)). The corresponding result is not always valid for Pareto optimal solutions. A counterexample can be found in Steuer (1986, p. 179). A result regarding the generation of the weakly Pareto optimal set of a convex problem as a union of such Pareto optimal sets where subsets of the objective functions are used, is proved in Lowe et al. (1984).

An overview is presented of duality theory for linear and nonlinear cases in Nakayama (1985c). Duality theory for nonlinear multiobjective optimization problems is also presented, for example, in Bitran (1981), Göpfert (1986), Luc (1984b, 1987, 1989), Nakayama (1984, 1985b, 1996), Singh et al. (1996) and Weir (1987); for convex problems in Jahn (1983) and Martínez-Legaz and Singer (1987); for more general convex-like problems in Das and Nanda (1997), Preda (1992, 1996) and Wang and Li (1992); for nonconvex problems in Luc and Jahn (1992); and for nonconvex nondifferentiable problems in Preda and Stancu-Minasian (1997). Some regularity results for multiobjective optimization problems are presented in Martein (1989). On the other hand, duality theory designed for a decision maker determining preferred solutions in convex multiobjective optimization problems is derived in Tarvainen (1996).

Finally, we state that an excellent account of stability and duality in multiobjective optimization can be found in Sawaragi et al. (1985). More than a third of the contents of the monograph addresses these topics.

Part II

METHODS

1. INTRODUCTION

Generating Pareto optimal solutions plays an important role in multiobjective optimization, and mathematically the problem is considered to be solved when the Pareto optimal set is found. The term *vector optimization* is sometimes used to denote the problem of identifying the Pareto optimal set. However, this is not always enough. We want to obtain only one solution. This means that we must find a way to put the Pareto optimal solutions in a complete order. This is why we need a decision maker and her or his preference structure. Here in Part II, we present several methods for solving multiobjective optimization problems. Usually, this means finding the Pareto optimal solution that best satisfies the decision maker.

We are not here going to interfere with the formulation of a real-life phenomenon as a mathematically well-defined problem. We merely stress that a proper formulation is important. Let us emphasize that in real-life problems inaccuracy in some form is often present. Remember that we exclude the handling of stochastic or fuzzy problems in this context. Even when the problems are modelled in a deterministic form, restricting the treatment to Pareto optimal solutions only may be misleading. For example, forgetting or misspelling an objective function may affect the Pareto optimal set. If it is impossible to model the practical problem in an explicit and precise mathematical form, we cannot automatically leave non-Pareto optimal solutions out of consideration. For example, imprecision of the data, the measurement or the objective functions means that the Pareto optimal set available is only an approximation of the real one. Here we have a gap between theory and practice.

Several crucial issues to bear in mind in the formulation of problems are treated in Haimes (1985) and Nijkamp et al. (1988). Among these are risk assessment, sufficient representativeness of the objective functions and precision of information. In many complicated, practical cases it may be impossible to give a correct formulation to the problem before it is solved. This means that the modelling and the solution phases should not be undertaken separately, which is generally the case nowadays. In other words, the modelling phase may require interaction with the solution phase. A parallel idea of approaching the modelling phase by including the decision maker in the modelling is suggested in Brans (1996). The goal is to give more freedom to the decision maker and not to limit her or his way of thinking to a prespecified model and its concepts.

In most methods we are interested in the objective space instead of the decision variable space. One reason for this is that the dimension of the objective space is usually considerably smaller than the dimension of the decision variable space. Another reason is that decision makers are often more interested in the objective values. However, calculation still takes place in the decision variable space because we do not usually know the explicit form of the feasible objective region. In brief, decision makers usually handle objective values whereas mathematical programming takes place in the decision variable space.

In general, multiobjective optimization problems are solved by scalarization. The most important exceptions to this are MOLP problems, which are not to be dealt with here, where some simplex-based solution methods can find Pareto optimal extreme points or, in some cases, the whole Pareto optimal set. Another exception, which is presented here, is the multiobjective proximal bundle method for nondifferentiable problems. It is not based on scalarization in the traditional sense.

As mentioned in Part I, scalarization means converting the problem into a single or a family of single objective optimization problems with a real-valued objective function, termed the *scalarizing function*, depending possibly on some parameters. This enables the use of the theory and the methods of scalar optimization, that is, nonlinear programming. Of fundamental importance is that the optimal solutions of multiobjective optimization problems can be characterized as solutions of certain single objective optimization problems. Because scalarizing functions usually depend on certain auxiliary parameters, some numerical difficulties may appear if the single objective optimization problem has feasible solutions only with very few parameter values or it is not solvable with all the parameter values. Thus the seemingly promising idea of simplifying the problem into single objective optimizations has also its weaknesses. In what follows, we assume that solutions to scalarizing functions exist.

In Sawaragi et al. (1985), three requirements are set for a scalarizing function:

(1) It can cover any Pareto optimal solution.
(2) Every solution is Pareto optimal.

If the scalarizing function is based on aspiration levels, then, in addition

(3) Its solution is satisficing if the aspiration levels used are feasible.

Unfortunately, there is no scalarizing function that can satisfy all three requirements.

An important fact to keep in mind is that standard routines for single objective optimization problems can only find local optima. This is why only locally Pareto optimal solutions are usually obtained and handled when dealing with scalarizing functions. Global Pareto optimality can be guaranteed, for example, if the objective functions and the feasible region are convex (as stated in Theorem 2.2.3 of Part I) or quasiconvex and convex, respectively (see Theorem

2.2.4 of Part I). An alternative is to employ global single objective optimizers. In the following, however, the solutions are understood to be local, unless stated otherwise.

Another matter to consider is the possibility of the scalarizing function having several alternative optimal solutions. In this case, the objective vector produced depends on the solution chosen. This may affect in an uncontrolled way the direction in which the solution process proceeds. This fact has not been taken into account in most method developments. Ideas for handling alternative optima in MOLP problems are presented in Sarma and Merouani (1995).

There is a large variety of methods for accomplishing multiobjective optimization. None of them can be said to be generally superior to all the others. When selecting a solution method, the specific features of the problem to be solved must be taken into consideration. In addition, the opinions of the decision maker are important. It is not enough that the analyst simply prefers some method. It may happen that the decision maker cannot or does not want to use it. The decision maker may be busy or mathematically ignorant. One can say that selecting an appropriate multiobjective optimization method itself is a problem with multiple objectives! We shall return to the method selection problem in Section 1.3 of Part III.

Methods of multiobjective optimization can be classified in many ways according to different criteria. In Cohon (1985), they are categorized into two relatively distinct subsets: generating methods and preference-based methods. In generating methods, the set of Pareto optimal (or efficient) solutions is generated for the decision maker, who then chooses one of the alternatives. In preference-based methods, the preferences of the decision maker are taken into consideration as the solution process goes on, and the solution that best satisfies the decision maker's preferences is selected.

Rosenthal (1985) suggests three classes of solution methods: partial generation of the Pareto optimal set, explicit value function maximization and interactive implicit value function maximization. In Carmichael (1981), methods are classified according to whether a composite single objective function, a single objective function with constraints, or many single objective functions are the basis for the approach. One more rough division could be made into interactive and noninteractive methods. These classes can be further divided in many ways.

Here we apply the classification presented in Hwang and Masud (1979). This classification is followed, for instance, in Buchanan (1986), Hwang et al. (1980) and Lieberman (1991a, b). Hwang and Masud classify the methods according to the participation of the decision maker in the solution process. The classes are:

1) methods where no articulation of preference information is used (*no-preference methods*),
2) methods where a posteriori articulation of preference information is used (*a posteriori methods*),

3) methods where a priori articulation of preference information is used (*a priori methods*), and
4) methods where progressive articulation of preference information is used (*interactive methods*).

For short, the names in the parentheses are used in the following.

However, no classification can be complete, as demonstrated, for example, in Despontin et al. (1983). Thus, one must bear in mind that the classifications are not absolute. Overlapping and combinations of classes are possible and some methods can be considered to belong to more than one class. The presented grouping is for guidance only.

In addition to the role of the decision maker we consider an alternative way of classification into ad hoc and non ad hoc methods. This division, suggested by Steuer and Gardiner (1991), is mainly intended for interactive methods, but can be applied to some other methods as well. It is based on the existence of an underlying value function. The common feature of *ad hoc* methods is that even if one knew the decision maker's value function, one would not exactly know how to respond to the questions posed by the algorithm. On the other hand, in *non ad hoc* methods the responses can be determined or at least confidently simulated if the decision maker's value function is known.

It should be pointed out that several concepts and assumptions underlying methods and solution processes can be questioned. For example, ten myths of multiobjective optimization and decision making are discussed and called into question in Michalowski (1997) and Zionts (1997a, b). Among them are concepts of well-defined decisions, isolated decision makers, optimal solutions, the value of Pareto optimal solutions, value functions, static decisions, preference of sophistication, mathematical convergence and technical assumptions. Here we do not go into details of these myths but refer to the presentations mentioned. A further aspect concerns the relative importance of objective functions. As emphasized in Roy and Mousseau (1996), such a notion is more complex than is usually recognized.

Before presenting the methods, we mention several references for further information. In Hwang and Masud (1979), a large number of methods is presented and illustrated by solving numerical examples in detail. A similar but shortened presentation is given in Hwang et al. (1980). The detailed solution process descriptions are intended to help in selecting solution methods.

Extensive surveys of concepts and methods for multiobjective optimization are provided in the monographs Chankong and Haimes (1983b) and Steuer (1986). Similar matters are studied briefly in Buchanan (1986), Chankong and Haimes (1983a), Chankong et al. (1985), Dyer and Sarin (1981), Rosenthal (1985), Steuer (1989b), Steuer and Gardiner (1990), Stewart (1992) and Vanderpooten (1990). An overview is given in Evans (1984) and several methods are also presented in Cohon (1985) and Osyczka (1984).

A set of methods developed up to the year 1973 for both multiattribute decision analysis and multiobjective optimization is collected in MacCrimmon

(1973). A wide collection of methods available (up to the year 1983) is assembled also in Despontin et al. (1983). Almost 100 methods for both multiobjective and multiattribute cases are included.

As far as different nationalities are concerned, overviews of multiobjective optimization methods in the former Soviet Union are presented in Lieberman (1991a, b) and of theory and applications in China in Hu (1990). Nine multiobjective optimization methods developed in Germany are briefly introduced in Ester and Holzmüller (1986).

A great number of interactive multiobjective optimization methods is collected in Shin and Ravindran (1991) and Vanderpooten and Vincke (1989). Interactive methods are also presented in Narula and Weistroffer (1989a) and White (1983b). Information about applications of the methods is also reported. Some literature on interactive multiobjective optimization between the years 1965 and 1988 is gathered in Aksoy (1990). A set of scalarizing functions is outlined in Wierzbicki (1986b) with special attention to whether weakly, properly or Pareto optimal solutions are produced.

As to different problem types, an overview of methods for MOLP problems can be found in Zionts (1980, 1989). Methods for hierarchical multiobjective optimization problems are reviewed in Haimes and Li (1988). Such methods are needed in large-scale problems. A wide survey on the literature of hierarchical multiobjective analysis is also provided.

Methods with applications to large-scale systems and industry are presented in the monographs Haimes et al. (1990) and Tabucanon (1988), respectively. Several groups of methods applicable to computer-aided design systems are presented briefly in Eiduks (1983). Methods for applications in structural optimization are reported in Eschenauer (1987), Jendo (1986), Koski and Silvennoinen (1987) and Osyczka and Koski (1989). The collections of papers edited by Eschenauer et al. (1990a) and Stadler (1988a) contain mainly applications in engineering.

In the following, we present several methods (in four classes) for multiobjective optimization. Some of them will be described in more detail and some only briefly mentioned. It must be kept in mind that the existing methodology is very wide. We do not intend to cover every existing method but to introduce several philosophies and ways of approaching multiobjective optimization problem solving. Where possible we try to link references to some of the applications and extensions available in the literature with the methods presented here. The description of each method ends with concluding remarks by the author taking up important aspects of the method. Unless stated otherwise, we assume that we solve problem (2.1.1) defined in Part I.

In connection with methods, a mention is made only of such implementations as have been made available to the author for testing purposes. By a *user* we mean either a decision maker or an analyst who uses the solution program. If the user is a decision maker, it is usually assumed that the problem has been formulated earlier (and perhaps loaded in the system) so that the decision maker can concentrate on the actual solution process.

2. NO-PREFERENCE METHODS

In no-preference methods, where the opinions of the decision maker are not taken into consideration, the multiobjective optimization problem is solved using some relatively simple method and the solution obtained is presented to the decision maker. The decision maker may either accept or reject the solution. It seems quite unlikely that the solution best satisfying the decision maker could be found with these methods. That is why no-preference methods are suitable for situations where the decision maker does not have any special expectations of the solution and (s)he is satisfied simply with some optimal solution. The working order here is: 1) analyst, 2) none.

As examples of this class we present the method of the global criterion and the multiobjective proximal bundle method.

2.1. Method of the Global Criterion

The method of the global criterion is also sometimes called *compromise programming* (see Yu (1973) and Zeleny (1973)). In this method, the distance between some reference point and the feasible objective region is minimized. The analyst has to select the reference point and the metric for measuring the distances. All the objective functions are thought to be equally important.

2.1.1. Different Metrics

Here we examine the method where the ideal objective vector is used as a reference point and L_p-metrics are used for measuring. In this case, the *L_p-problem* to be solved is

$$\begin{aligned} &\text{minimize} && \left(\sum_{i=1}^{k} \left| f_i(\mathbf{x}) - z_i^\star \right|^p \right)^{1/p} \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{2.1.1}$$

From the definition of the ideal objective vector $\mathbf{z}^\star$ we know that $f_i(\mathbf{x}) \geq z_i^\star$ for all $i = 1, \ldots, k$ and all $\mathbf{x} \in S$. This is why no absolute values are needed if we know the global ideal objective vector. If the global ideal objective vector

K. Miettinen, *Nonlinear Multiobjective Optimization*

is not known, the method does not necessarily work as it should. In order to emphasize this fact, we keep the absolute value signs in the notations when introducing the method.

If the ideal objective vector is replaced by some other vector, it must be selected carefully. Pessimistic reference points must be avoided since the method cannot find solutions better than the reference point.

The exponent $1/p$ may be dropped. Problems with or without the exponent $1/p$ are equivalent for $1 \leq p < \infty$, since L_p-problem (2.1.1) is an increasing function of the corresponding problem without the exponent.

If $p = \infty$, the metric is also called a *Tchebycheff metric* and the $L_\infty-$ or the *Tchebycheff problem* is of the form

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} \left[|f_i(\mathbf{x}) - z_i^\star| \right] \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{2.1.2}$$

Notice that problem (2.1.2) is nondifferentiable even in the absence of absolute values. In this case, it can, however, be transformed into a differentiable form if the objective and the constraint functions are differentiable. Then, instead of problem (2.1.2), the problem

$$\begin{aligned} &\text{minimize} && \alpha \\ &\text{subject to} && \alpha \geq f_i(\mathbf{x}) - z_i^\star \quad \text{for all} \quad i = 1, \ldots, k, \\ & && \mathbf{x} \in S, \end{aligned}$$

is solved, where both $\mathbf{x} \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$ are variables.

The solution obtained depends greatly on the value chosen for p. Widely used choices are $p = 1, 2$ or ∞. In Figure 2.1.1, the contours of these three different metrics are shown. The black point is the ideal objective vector and the bold line represents the Pareto optimal set. It is worth noticing that if the original problem is linear, the choice $p = 1$ maintains the linearity. As the value of p increases, the nonlinear minimization problem becomes more difficult and badly conditioned to solve.

For linear problems, the solutions obtained by the L_p-problems where $1 < p < \infty$ are situated between the solutions obtained by the L_1- and L_∞-problems. It is illustrated in Zeleny (1973) that this set of solutions is a part of the Pareto optimal set, but only a substantially small part.

Instead of the terms $|f_i(\mathbf{x}) - z_i^\star|$, denominators may be added to problems (2.1.1) and (2.1.2) to normalize the components, that is, to use the terms $|f_i(\mathbf{x}) - z_i^\star| \,/\, |z_i^\star|$ instead. Some other denominators, like $|z_i^{\text{nad}} - z_i^\star|$, can also be used. The reason for employing denominators is that sometimes it is worthwhile to use relative distances in the calculations. For example, using the components of $\mathbf{z}^\star$ forms the contour of the metric to reflect better the location of the ideal objective vector. Naturally, the denominators $z_i^\star$ cannot be used if some of them equals zero.

The objective functions may also be normalized by

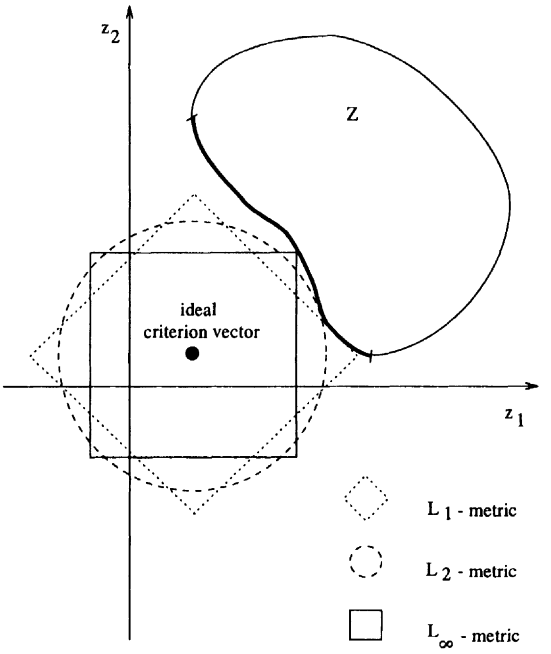

Figure 2.1.1. Different metrics.

$$\bar{f}_i(\mathbf{x}) = \frac{f_i(\mathbf{x}) - z_i^\star}{\max_{\mathbf{x}\in S} f_i(\mathbf{x}) - z_i^\star} \tag{2.1.3}$$

before the distance is minimized. In this case, the range of the new objective functions is $[0, 1]$. This normalizing is possible only if the objectives are bounded. However, it is usually better to employ the ranges of the Pareto optimal set and replace the max term by the component of the approximated nadir objective vector z_i^{nad} in (2.1.3).

A variation of the Tchebycheff problem is suggested in Osyczka (1989a, 1992), where the problem to be solved is

$$\begin{aligned} &\text{minimize} \quad \max_{i=1,\ldots,k} \left[\max \left[\left| \frac{f_i(\mathbf{x}) - z_i^\star}{z_i^\star} \right|, \left| \frac{f_i(\mathbf{x}) - z_i^\star}{f_i(\mathbf{x})} \right| \right] \right] \\ &\text{subject to} \quad \mathbf{x} \in S. \end{aligned} \tag{2.1.4}$$

2.1.2. Theoretical Results

Next, we present some theoretical results concerning the method of the global criterion. We assume that we know the global ideal objective vector and can, thus, leave the absolute values.

Theorem 2.1.1. The solution of L_p-problem (2.1.1) (where $1 \leq p < \infty$) is Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a solution of problem (2.1.1) with $1 \leq p < \infty$. Let us suppose that $\mathbf{x}^*$ is not Pareto optimal. Then there exists a point $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$ for at least one j. Now $(f_i(\mathbf{x}) - z_i^\star)^p \leq (f_i(\mathbf{x}^*) - z_i^\star)^p$ for all i and $(f_j(\mathbf{x}) - z_j^\star)^p < (f_j(\mathbf{x}^*) - z_j^\star)^p$. From this we obtain

$$\sum_{i=1}^{k} \left(f_i(\mathbf{x}) - z_i^\star\right)^p < \sum_{i=1}^{k} \left(f_i(\mathbf{x}^*) - z_i^\star\right)^p.$$

When both sides of the inequality are raised into the power $1/p$ we have a contradiction to the assumption that $\mathbf{x}^*$ is a solution of problem (2.1.1). This completes the proof. □

Yu has pointed out in Yu (1973) that if Z is a convex set, then for $1 < p < \infty$ the solution of problem (2.1.1) is unique.

Theorem 2.1.2. The solution of Tchebycheff problem (2.1.2) is weakly Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a solution of problem (2.1.2). Let us suppose that $\mathbf{x}^*$ is not weakly Pareto optimal. In this case, there exists a point $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$. It means that, $f_i(\mathbf{x}) - z_i^\star < f_i(\mathbf{x}^*) - z_i^\star$ for all i. Thus, $\mathbf{x}^*$ cannot be a solution of problem (2.1.2). Here we have a contradiction which completes the proof. □

Theorem 2.1.3. Tchebycheff problem (2.1.2) has at least one Pareto optimal solution.

Proof. Let us suppose that none of the optimal solutions of problem (2.1.2) is Pareto optimal. Let $\mathbf{x}^* \in S$ be an optimal solution of problem (2.1.2). Since we assume that it is not Pareto optimal, there must exist a solution $\mathbf{x} \in S$ which is not optimal for problem (2.1.2) but for which $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$ for at least one j.

We have now $f_i(\mathbf{x}) - z_i^\star \leq f_i(\mathbf{x}^*) - z_i^\star$ for all i with the strict inequality holding for at least one index j, and further $\max_i[f_i(\mathbf{x}) - z_i^\star] \leq \max_i[f_i(\mathbf{x}^*) - z_i^\star]$. Because $\mathbf{x}^*$ is an optimal solution of problem (2.1.2), $\mathbf{x}$ has to be an optimal solution, as well. This contradiction completes the proof. □

Corollary 2.1.4. If Tchebycheff problem (2.1.2) has a unique solution, it is Pareto optimal.

A linear numerical application example of the method is given in Hwang and Masud (1979, pp. 23–29). Sufficient conditions for the solution of an L_p-problem

to be stable with respect to changes in the feasible region S are presented in Jurkiewicz (1983). Reference points more general than the ideal objective vector guaranteeing Pareto optimal results are handled in Skulimowski (1996).

2.1.3. Concluding Remarks

The method of the global criterion is a simple method to use if the aim is simply to obtain a solution where no special hopes are set. The properties of the metrics imply that if the objective functions are not normalized in any way, then an objective function whose ideal objective value is situated nearer the feasible objective region receives more importance.

The solution obtained with the L_p-metric ($1 \leq p < \infty$) is guaranteed to be Pareto optimal. If the Tchebycheff metric is used, the solution may be weakly Pareto optimal. In the latter case, for instance, problem (2.10.2) of Part I can be used to produce Pareto optimal solutions. It is up to the analyst to select an appropriate metric.

2.2. Multiobjective Proximal Bundle Method

The multiobjective proximal bundle (MPB) method is an extension of single-objective bundle-based methods of nondifferentiable optimization into the multiobjective case. It is derived in Mäkelä (1993) and Miettinen and Mäkelä (1995, 1996a) according to the ideas of Kiwiel (1984, 1985a, b) and Wang (1989). The underlying proximal bundle method, presented in Kiwiel (1990), is an advanced version of the bundle family for convex, unconstrained nondifferentiable single objective optimization. It is generalized for nonconvex and constrained problems in Mäkelä and Neittaanmäki (1992, pp. 112–137).

The idea of the MPB method, in brief, is to move in a direction where the values of all the objective functions improve simultaneously. Here we describe features of the MPB method from an implementational viewpoint. For details see Mäkelä (1993).

2.2.1. Introduction

The MPB method is capable of solving problems with nonlinear (possibly nondifferentiable) functions. It is assumed that all the objective and the constraint functions are locally Lipschitzian.

The MPB method is not like the other scalarization methods. Ordinary scalarization methods transform the problem into one with a single objective function. This new problem can then be solved with any appropriate method for nonlinear programming. In the MPB method the scalarizing function lies inside a special (nondifferentiable) optimizer, which is why its philosophy is so

different from that of the other methods described here, and why implementational aspects that can be forgotten with other methods have to be touched on.

Let us now suppose that the feasible region is of the form

$$S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}.$$

As mentioned above, the MPB method is not directly based on employing any scalarizing function. Some kind of scalarization is, however, needed in deriving the minimization method for all the objective functions. Theoretically, we utilize an unconstrained *improvement function* $H\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ defined by

$$H(\mathbf{x}^1, \mathbf{x}^2) = \max[f_i(\mathbf{x}^1) - f_i(\mathbf{x}^2),\ g_l(\mathbf{x}^1) \mid i = 1, \ldots, k,\ l = 1, \ldots, m].$$

Let us first prove a result about the optimal solutions of improvement functions. The sufficient condition necessitates the Slater constraint qualification (Definition 3.2.13 in Part I).

Theorem 2.2.1. A necessary condition for a point $\mathbf{x}^* \in \mathbf{R}^n$ to be weakly Pareto optimal is that the improvement function $H(\mathbf{x}, \mathbf{x}^*)$ attains its minimum at $\mathbf{x}^*$. If the multiobjective optimization problem is convex and the Slater constraint qualification is satisfied, then the condition above is also sufficient.

Proof. The necessity follows immediately from the proof of Theorem 3.2.5 (and Corollary 3.2.7) in Section 3.2 of Part I.

As to the sufficiency component, let the assumptions stated be valid and let $\mathbf{x}^* \in \mathbf{R}^n$ be a minimal solution of $H(\mathbf{x}, \mathbf{x}^*)$. Let us assume that $\mathbf{x}^*$ is not weakly Pareto optimal. Then, there exists some $\hat{\mathbf{x}} \in \mathbf{R}^n$ such that $g_j(\hat{\mathbf{x}}) \leq 0$ for all $j = 1, \ldots, m$ and $f_i(\hat{\mathbf{x}}) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$. If $g_j(\hat{\mathbf{x}}) < 0$ for all $j = 1, \ldots, m$, then

$$H(\hat{\mathbf{x}}, \mathbf{x}^*) < 0 = H(\mathbf{x}^*, \mathbf{x}^*),$$

which contradicts the assumption that $H(\mathbf{x}, \mathbf{x}^*)$ attains its minimum at $\mathbf{x}^*$.

Otherwise, that is, if $g_j(\hat{\mathbf{x}}) = 0$ for some index j, it follows from the Slater constraint qualification that there exists some $\bar{\mathbf{x}} \in \mathbf{R}^n$ such that $g_j(\bar{\mathbf{x}}) < 0$ for all $j = 1, \ldots, m$. If $f_i(\bar{\mathbf{x}}) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$, then

$$H(\bar{\mathbf{x}}, \mathbf{x}^*) < 0 = H(\mathbf{x}^*, \mathbf{x}^*)$$

and we have a contradiction with $\mathbf{x}^*$ minimizing H.

Otherwise, we define $I_0 \subset \{1, \ldots, k\}$ such that $f_i(\bar{\mathbf{x}}) \geq f_i(\mathbf{x}^*) > f_i(\hat{\mathbf{x}})$ for all $i \in I_0$. Let us denote $\mathbf{y} = \lambda\bar{\mathbf{x}} + (1 - \lambda)\hat{\mathbf{x}}$ for $0 < \lambda < 1$. Then the convexity of the constraint and the objective functions implies that for all $0 < \lambda < 1$

$$g_j(\mathbf{y}) \leq \lambda g_j(\bar{\mathbf{x}}) + (1 - \lambda) g_j(\hat{\mathbf{x}}) < 0 \tag{2.2.1}$$

for all $j = 1, \ldots, m$, and

$$f_i(\mathbf{y}) \leq \lambda f_i(\bar{\mathbf{x}}) + (1-\lambda) f_i(\hat{\mathbf{x}}) < \lambda f_i(\mathbf{x}^*) + (1-\lambda) f_i(\mathbf{x}^*) = f_i(\mathbf{x}^*)$$

for all $i \in \{1,\ldots,k\} \setminus I_0$. If I_0 is nonempty, we choose

$$\lambda = \frac{\min\left[f_i(\mathbf{x}^*) - f_i(\hat{\mathbf{x}}) \mid i \in I_0\right]}{\max\left[f_i(\bar{\mathbf{x}}) - f_i(\hat{\mathbf{x}}) \mid i \in I_0\right]} - \varepsilon,$$

where $\varepsilon > 0$ is small enough so that $0 < \lambda < 1$. By the convexity of the objective functions, we have for all $i \in I_0$

(2.2.2)

$$\begin{aligned} f_i(\mathbf{y}) &\leq \frac{f_i(\mathbf{x}^*) - f_i(\hat{\mathbf{x}})}{f_i(\bar{\mathbf{x}}) - f_i(\hat{\mathbf{x}})} f_i(\bar{\mathbf{x}}) + \left(1 - \frac{f_i(\mathbf{x}^*) - f_i(\hat{\mathbf{x}})}{f_i(\bar{\mathbf{x}}) - f_i(\hat{\mathbf{x}})}\right) f_i(\hat{\mathbf{x}}) - \varepsilon(f_i(\bar{\mathbf{x}}) - f_i(\hat{\mathbf{x}})) \\ &= f_i(\mathbf{x}^*) - \varepsilon(f_i(\bar{\mathbf{x}}) - f_i(\hat{\mathbf{x}})) < f_i(\mathbf{x}^*). \end{aligned}$$

Then, combining the results (2.2.1) and (2.2.2), we obtain

$$H(\mathbf{y}, \mathbf{x}^*) < 0 = H(\mathbf{x}^*, \mathbf{x}^*),$$

which is again a contradiction with the condition that $\mathbf{x}^*$ minimized H. Thus, $\mathbf{x}^*$ is weakly Pareto optimal. □

2.2.2. MPB Algorithm

In the following, we take a look at the MPB method. We do not describe the method completely but present its idea roughly. The reason is that the structure of the method is highly connected to the underlying nondifferentiable proximal bundle method.

In the MPB method, the solution is looked for iteratively, until some stopping criterion is fulfilled. The iteration counter h refers to the inner iterations of the MPB method. Let $\mathbf{x}^h$ be the current approximation to the solution of the multiobjective optimization problem at the iteration h. Then, by Theorem 2.2.1, we seek for the search direction $\mathbf{d}^h$ as a solution of the unconstrained optimization problem

(2.2.3)

$$\begin{aligned} &\text{minimize} && H(\mathbf{x}^h + \mathbf{d}, \mathbf{x}^h) \\ &\text{subject to} && \mathbf{d} \in \mathbf{R}^n. \end{aligned}$$

Since problem (2.2.3) is still nondifferentiable, we must approximate it somehow.

Let us assume for a moment that the problem is convex. We suppose that, at the iteration h in addition to the iteration point $\mathbf{x}^h$, we have some auxiliary points $\mathbf{y}^j \in \mathbf{R}^n$ from past iterations and subgradients $\boldsymbol{\xi}_{f_i}^j \in \partial f_i(\mathbf{y}^j)$ for $j \in J^h$, $i = 1,\ldots,k$, and $\boldsymbol{\xi}_{g_l}^j \in \partial g_l(\mathbf{y}^j)$ for $j \in J^h$, $l = 1,\ldots,m$, where J^h is a nonempty subset of $\{1,\ldots,h\}$. We linearize the objective and the constraint functions at the point $\mathbf{y}^j$ and denote

$$\bar{f}_{i,j}(\mathbf{x}) = f_i(\mathbf{y}^j) + (\boldsymbol{\xi}_{f_i}^j)^T(\mathbf{x} - \mathbf{y}^j) \quad \text{for all } i = 1,\ldots,k,\ j \in J^h \quad \text{and}$$
$$\bar{g}_{l,j}(\mathbf{x}) = g_l(\mathbf{y}^j) + (\boldsymbol{\xi}_{g_l}^j)^T(\mathbf{x} - \mathbf{y}^j) \quad \text{for all } l = 1,\ldots,m,\ j \in J^h.$$

We can now define a convex piecewise linear approximation to the improvement function by

$$\hat{H}^h(\mathbf{x}) = \max\left[\bar{f}_{i,j}(\mathbf{x}) - f_i(\mathbf{x}^h),\ \bar{g}_{l,j}(\mathbf{x}) \mid i = 1,\ldots,k,\ l = 1,\ldots,m,\ j \in J^h\right]$$

and we get an approximation to (2.2.3) by

$$\begin{aligned} &\text{minimize} && \hat{H}^h(\mathbf{x}^h + \mathbf{d}) + \tfrac{1}{2}u^h\|\mathbf{d}\|^2 \\ &\text{subject to} && \mathbf{d} \in \mathbf{R}^n, \end{aligned} \tag{2.2.4}$$

where $u^h > 0$ is an inner parameter to be updated automatically. The penalty term $\frac{1}{2}u^h\|\mathbf{d}\|^2$ is added to guarantee that there exists a solution to problem (2.2.4) and to keep the approximation local enough.

Notice that (2.2.4) is still a nondifferentiable problem, but due to its min-max-nature it is equivalent to the following differentiable quadratic problem with $\mathbf{d} \in \mathbf{R}^n$ and $v \in \mathbf{R}$ as variables:

$$\begin{aligned} &\text{minimize} && v + \tfrac{1}{2}u^h\|\mathbf{d}\|^2 \\ &\text{subject to} && v \geq -\alpha_{f_i,j}^h + (\boldsymbol{\xi}_{f_i}^j)^T\mathbf{d},\ i = 1,\ldots,k,\ j \in J^h \\ & && v \geq -\alpha_{g_l,j}^h + (\boldsymbol{\xi}_{g_l}^j)^T\mathbf{d},\ l = 1,\ldots,m,\ j \in J^h, \end{aligned} \tag{2.2.5}$$

where

$$\alpha_{f_i,j}^h = f_i(\mathbf{x}^h) - \bar{f}_{i,j}(\mathbf{x}^h),\ i = 1,\ldots,k,\ j \in J^h \quad \text{and}$$
$$\alpha_{g_l,j}^h = -\bar{g}_{l,j}(\mathbf{x}^h),\ l = 1,\ldots,m,\ j \in J^h$$

are so-called linearization errors.

In the nonconvex case, we replace the linearization errors by so-called subgradient locality measures:

$$\beta_{f_i,j}^h = \max\left[|\alpha_{f_i,j}^h|, \gamma_{f_i}\|\mathbf{x}^h - \mathbf{y}^j\|^2\right]$$
$$\beta_{g_l,j}^h = \max\left[|\alpha_{g_l,j}^h|, \gamma_{g_l}\|\mathbf{x}^h - \mathbf{y}^j\|^2\right],$$

where $\gamma_{f_i} \geq 0$ for $i = 1,\ldots,k$ and $\gamma_{g_l} \geq 0$ for $l = 1,\ldots,m$ are so-called distance measure parameters ($\gamma_{f_i} = 0$ if f_i is convex and $\gamma_{g_l} = 0$ if g_l is convex).

Let $(\mathbf{d}^h, v^h)$ be a solution of problem (2.2.5). Then the two-point line search strategy is carried out. It detects discontinuities in the gradients of the objective functions. Roughly speaking, we try to find a step-size $0 < t^h \leq 1$ such that $H(\mathbf{x}^h + t^h\mathbf{d}^h, \mathbf{x}^h)$ is minimal when $\mathbf{x}^h + t^h\mathbf{d}^h \in S$.

A line search algorithm (in Mäkelä and Neittaanmäki (1992, pp. 126–130)) is used to produce the step-sizes. The iteration is terminated when a predetermined accuracy is reached. The subgradient aggregation strategy due to Kiwiel (1985c) is used to bound the storage requirements (i.e., the size of the

index set J^h) and a modification of the weight-updating algorithm described in Kiwiel (1990) is used to update the weight u^h. For details, see Miettinen and Mäkelä (1995, 1998a).

This is roughly the MPB method. Next, some words about optimality are in order.

2.2.3. Theoretical Results

According to Theorem 2.2.1 we, on the one hand, know that minimizing an improvement function produces weakly Pareto optimal solutions. On the other hand, any weakly Pareto optimal solution of a convex problem can be found under minor conditions. While we do not optimize the improvement function but its approximation, the optimality results of the MPB method are somewhat different. Here we only present some results without proofs, since giving these would necessitate explicit expression of the MPB algorithm.

Theorem 2.2.2. Let the multiobjective optimization problem be convex and the Slater constraint qualification be satisfied. If the MPB method stops with a finite number of iterations, then the solution is weakly Pareto optimal. On the other hand, any accumulation point of the infinite sequence of solutions generated by the MPB method is weakly Pareto optimal.

Proof. See Kiwiel (1985a) or Wang (1989).

If the convexity assumption is not satisfied, we obtain somewhat weaker results about substationary points (See Definition 3.2.6 of Part I). This result involves upper semidifferentiable functions (see Definition 2.1.15 of Part I).

Theorem 2.2.3. Let the objective and the constraint functions of the multiobjective optimization problem be upper semidifferentiable at every $\mathbf{x} \in S$. If the MPB method stops with a finite number of iterations, then the solution is a substationary point. On the other hand, any accumulation point of an infinite sequence of solutions generated by the MPB method is a substationary point.

Proof. See Wang (1989) and references therein.

Note that only the substationarity of the solutions of the MPB routine is guaranteed for general multiobjective optimization problems.

2.2.4. Concluding Remarks

The MPB method can be used as a method where no opinions of the decision maker are sought. In this case, we must select the starting point so that it is not (weakly) Pareto optimal but that every component of the objective vector can

be improved. The method can also handle other than nonlinear constraints, but they have not been included here for the sake of the clarity of the presentation.

The MPB routine can also be used as a black-box optimized within interactive multiobjective optimization methods. This is the case with the vector version of NIMBUS (see Section 5.12).

The accuracy of the computation in the MPB method is an interesting matter. Accuracy can be considered in a more extensive meaning as a separating factor between ordinary scalarizing functions and inner scalarizing function, as in the MPB method. If some ordinary scalarizing function is employed, then it is the accuracy of that additional function that can be followed along with the solution process. It may happen that when the accuracy of the scalarizing function has reached the desired level, the values of the actual objective functions could still change considerably.

Many scalarizing functions have positive features whose importance is not to be underestimated, such as producing only Pareto optimal solutions. However, employing some scalarizing function usually brings along extra parameters and the difficulty of specifying their values. This causes additional stability concern. To put it briefly, scalarizing functions add extra characteristics to the problem.

Scalarization cannot completely be avoided even in the MPB routine. However, the scalarization is carried out under the surface, invisible to the user. Whatever additional parameters or phases are needed, they cannot be seen and the user does not have to be bothered with them. The weakness of the MPB routine is that the Pareto optimality of the solutions obtained cannot be guaranteed. In theory, only the substationarity of the solutions is certain. In practice, it is, however, very likely that the solutions are at least weakly Pareto optimal. As a matter of fact, in the numerical experiments performed, the final solutions obtained have usually proved to be Pareto optimal at the final testing.

For problems with nondifferentiable functions the MPB routine represents an efficient proximal bundle-based solution approach. The implementation of the MPB routine (called MPBNGC) is described in Mäkelä (1993). It calls a quadratic solver derived in Kiwiel (1986).

3. A POSTERIORI METHODS

A posteriori methods could also be called *methods for generating Pareto optimal solutions*. After the Pareto optimal set (or a part of it) has been generated, it is presented to the decision maker, who selects the most preferred among the alternatives. The inconveniences here are that the generation process is usually computationally expensive and sometimes in part, at least, difficult. On the other hand, it is hard for the decision maker to select from a large set of alternatives. One more important question is how to present or display the alternatives to the decision maker in an effective way. The working order in these methods is: 1) analyst, 2) decision maker.

If there are only two objective functions, the Pareto optimal set can be generated parametrically (see, for example, Benson (1979b) and Gass and Saaty (1955)). When there are more objectives, the problem becomes more complicated.

Let us briefly mention that in solving MOLP problems the methods can be divided into two subclasses. In the first are the methods that can find all the Pareto optimal solutions and in the second are the methods that can find only all the Pareto optimal extreme solutions. In the latter case, edges connecting Pareto optimal extreme points may be Pareto optimal or not. In nonlinear problems, the distinction lies between convex and nonconvex problems. In other words, some methods can only generate Pareto optimal solutions of convex problems.

The methods presented in detail here are called *basic methods*, since they are used frequently in practical problems, and they are also used as elements of more developed methods. Basic methods are the weighting method and the ε-constraint method. After them, we give a limited overview of a method combining features of both the weighting and the ε-constraint method. Then we introduce two more basic methods. The method of weighted metrics is a weighted extension of the method of the global criterion. It is followed by the handling of achievement scalarizing functions. Finally, some other methods in this class are briefly mentioned.

K. Miettinen, *Nonlinear Multiobjective Optimization*

3.1. Weighting Method

In the weighting method, presented, for example, in Gass and Saaty (1955) and Zadeh (1963), the idea is to associate each objective function with a weighting coefficient and minimize the weighted sum of the objectives. In this way, the multiple objective functions are transformed into a single objective function. We suppose that the weighting coefficients w_i are real numbers such that $w_i \geq 0$ for all $i = 1, \ldots, k$. It is also usually supposed that the weights are normalized, that is, $\sum_{i=1}^{k} w_i = 1$. To be more exact, the multiobjective optimization problem is modified into the following problem, to be called a *weighting problem*:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{k} w_i f_i(\mathbf{x}) \\ &\text{subject to} && \mathbf{x} \in S, \end{aligned} \tag{3.1.1}$$

where $w_i \geq 0$ for all $i = 1, \ldots, k$ and $\sum_{i=1}^{k} w_i = 1$.

3.1.1. Theoretical Results

In the following, several theoretical results are presented concerning the weighting method.

Theorem 3.1.1. The solution of weighting problem (3.1.1) is weakly Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a solution of the weighting problem. Let us suppose that it is not weakly Pareto optimal. In this case, there exists a solution $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$. According to the assumptions set to the weighting coefficients, $w_j > 0$ for at least one j. Thus we have $\sum_{i=1}^{k} w_i f_i(\mathbf{x}) < \sum_{i=1}^{k} w_i f_i(\mathbf{x}^*)$. This is a contradiction to the assumption that $\mathbf{x}^*$ is a solution of the weighting problem. Thus $\mathbf{x}^*$ is weakly Pareto optimal. □

Theorem 3.1.2. The solution of weighting problem (3.1.1) is Pareto optimal if the weighting coefficients are positive, that is $w_i > 0$ for all $i = 1, \ldots, k$.

Proof. Let $\mathbf{x}^* \in S$ be a solution of the weighting problem with positive weighting coefficients. Let us suppose that it is not Pareto optimal. This means that there exists a solution $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$ for at least one j. Since $w_i > 0$ for all $i = 1, \ldots, k$, we have $\sum_{i=1}^{k} w_i f_i(\mathbf{x}) < \sum_{i=1}^{k} w_i f_i(\mathbf{x}^*)$. This contradicts the assumption that $\mathbf{x}^*$ is a solution of the weighting problem and, thus, $\mathbf{x}^*$ must be Pareto optimal. □

Theorem 3.1.3. The unique solution of weighting problem (3.1.1) is Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a unique solution of the weighting problem. Let us suppose that it is not Pareto optimal. In this case, there exists a solution $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$ for at least one j. Because all the weighting coefficients w_i are nonnegative, we have $\sum_{i=1}^{k} w_i f_i(\mathbf{x}) \leq \sum_{i=1}^{k} w_i f_i(\mathbf{x}^*)$. On the other hand, the uniqueness of $\mathbf{x}^*$ means that $\sum_{i=1}^{k} w_i f_i(\mathbf{x}^*) < \sum_{i=1}^{k} w_i f_i(\hat{\mathbf{x}})$ for all $\hat{\mathbf{x}} \in S$. The two inequalities above are contradictory and, thus, $\mathbf{x}^*$ must be Pareto optimal. □

As Theorems 3.1.2 and 3.1.3 state, the solution of the weighting method is always Pareto optimal if the weighting coefficients are all positive or if the solution is unique, without any further assumptions. The weakness of the weighting method is that not all of the Pareto optimal solutions can be found unless the problem is convex. This feature can be relaxed to some extent by convexifying the nonconvex Pareto optimal set as suggested in Li (1996). The convexification is realized by raising the objective functions to a high enough power under certain assumptions. However, the main result is the following:

Theorem 3.1.4. Let the multiobjective optimization problem be convex. If $\mathbf{x}^* \in S$ is Pareto optimal, then there exists a weighting vector $\mathbf{w}$ ($w_i \geq 0$, $i = 1, \ldots, k$, $\sum_{i=1}^{k} w_i = 1$) such that $\mathbf{x}^*$ is a solution of weighting problem (3.1.1).

Proof. The proof is presented after Theorem 3.2.6.

According to Theorem 3.1.4 any Pareto optimal solution of a convex multiobjective optimization problem can be found by the weighting method. Note that the weighting vector is not necessarily unique. The contents of Theorem 3.1.4 is illustrated in Figure 3.1.1. On the left, every Pareto optimal solution along the bold line can be obtained by altering the weighting coefficients. On the right, it is not possible to obtain the Pareto optimal solutions in the 'hollow.'

An equivalent trigonometric formulation to the weighting problem with two objective functions is presented in Das and Dennis (1997). This formulation can be used in illustrating geometrically why not all the Pareto optimal solutions of nonconvex problems can be found.

Remark 3.1.5. According to Theorem 3.1.4, all the Pareto optimal solutions of MOLP problems can be found by the weighting method.

Let us have a look at linear cases for a while. In practice, Remark 3.1.5 is not quite true. The single objective optimization routines for linear problems

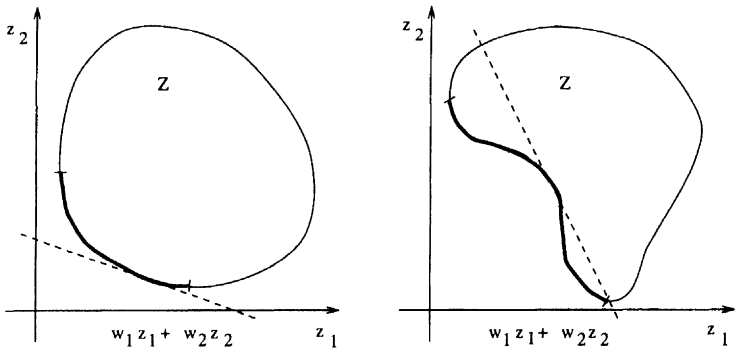

Figure 3.1.1. Weighting method with convex and nonconvex problems.

usually find only extreme point solutions. Thus, if some facet of the feasible region is Pareto optimal, then the infinity of Pareto optimal non-extreme points must be described in terms of linear combinations of the Pareto optimal extreme solutions. On the other hand, note that if two adjacent Pareto optimal extreme points for an MOLP problem are found, the edge connecting them is not necessarily Pareto optimal.

The conditions under which the whole Pareto optimal set can be generated by the weighting method with positive weighting coefficients are presented in Censor (1977). The solutions that it is possible to reach by the weighting method with positive weighting coefficients are characterized in Belkeziz and Pirlot (1991). Some generalized results are also given. More relations between nonnegative and positive weighting coefficients, convexity of S and Z, and Pareto optimality are studied in Lin (1976b).

If the weighting coefficients in the weighting method are all positive, we can say more about the solutions than that they are Pareto optimal. The following results concerning proper Pareto optimality were originally presented in Geoffrion (1968).

Theorem 3.1.6. The solution of weighting problem (3.1.1) is properly Pareto optimal if all the weighting coefficients are positive (sufficient condition).

Proof. Let $\mathbf{x}^* \in S$ be a solution of the weighting problem with positive weighting coefficients. In Theorem 3.1.2 we showed that the solution is Pareto optimal. We shall now show that $\mathbf{x}^*$ is properly Pareto optimal with $M = (k-1)\max_{i,j}(w_j/w_i)$.

Let us on the contrary suppose that $\mathbf{x}^*$ is not properly Pareto optimal. Then for some i (which we fix) and for $\mathbf{x} \in S$ such that $f_i(\mathbf{x}^*) > f_i(\mathbf{x})$ we have

$$f_i(\mathbf{x}^*) - f_i(\mathbf{x}) > M(f_j(\mathbf{x}) - f_j(\mathbf{x}^*))$$

for all j such that $f_j(\mathbf{x}^*) < f_j(\mathbf{x})$. We can now write

$$f_i(\mathbf{x}^*) - f_i(\mathbf{x}) > (k-1)\frac{w_j}{w_i}(f_j(\mathbf{x}) - f_j(\mathbf{x}^*)).$$

After multiplying both sides by $w_i/(k-1) > 0$, we get

$$\frac{w_i}{k-1}(f_i(\mathbf{x}^*) - f_i(\mathbf{x})) > w_j(f_j(\mathbf{x}) - f_j(\mathbf{x}^*)) \left(> 0 \geq w_l(f_l(\mathbf{x}) - f_l(\mathbf{x}^*))\right),$$

where l differs from the fixed index i and the indices j, which were specified earlier. After this reasoning we can sum over all $j \neq i$ and obtain

$$w_i(f_i(\mathbf{x}^*) - f_i(\mathbf{x})) > \sum_{\substack{j=1 \\ j\neq i}}^{k}(w_j(f_j(\mathbf{x}) - f_j(\mathbf{x}^*))),$$

which means

$$\sum_{j=1}^{k} w_j f_j(\mathbf{x}^*) > \sum_{j=1}^{k} w_j f_j(\mathbf{x}).$$

Here we have a contradiction to the assumption that $\mathbf{x}^*$ is a solution of the weighting problem. Thus, $\mathbf{x}^*$ has to be properly Pareto optimal. □

Theorem 3.1.7. If the multiobjective optimization problem is convex, then the condition in Theorem 3.1.6 is also necessary.

Proof. See Geoffrion (1968) or Chou et al. (1985).

Corollary 3.1.8. A necessary and sufficient condition for a point to be a properly Pareto optimal solution of an MOLP problem is that it is a solution of a weighting problem with all the weighting coefficients being positive.

The ratio of the weighting coefficients gives an upper bound to global trade-offs.

Theorem 3.1.9. Let $\mathbf{x}^*$ be a solution of weighting problem (3.1.1), when all the weighting coefficients w_i, $i = 1, \ldots, k$, are positive. Then

$$\Lambda_{ij}^{G}(\mathbf{x}^*) \leq \max_{\substack{i=1,\ldots,k \\ i\neq j}} \frac{w_j}{w_i}$$

for every $i, j = 1, \ldots, k$, $i \neq j$.

Proof. See Kaliszewski (1994, p. 93).

Some results concerning weak, proper and Pareto optimality of the solutions obtained by the weighting method are combined in Wierzbicki (1986b). Proper Pareto optimality and the weighting method are also discussed in Belkeziz and Pirlot (1991) and Luc (1995). The weighting method is used in Isermann (1974)

in proving that for linear multiobjective optimization problems all the Pareto optimal solutions are also properly Pareto optimal. More results concerning MOLP problems and the weighting method are assembled in Chankong and Haimes (1983a, b, pp. 153–159).

3.1.2. Applications and Extensions

As far as applications are concerned, the weighting method is used to generate Pareto optimal solutions in Sadek et al. (1988–89) in solving a problem of the optimal control of a damped beam. The weighting method is also applied in Weck and Förtsch (1988) to structural systems in the optimization of a spindle bearing system and in the optimization of a table, as well as in ReVelle (1988), where reductions in strategic nuclear weapons for the two superpowers are examined. Furthermore, Pareto optimal solutions are generated for an anti-lock brake system control problem by the weighting method in Athan and Papalambros (1997). However, no attention is paid to the possible nonconvexity of the problem. In addition, the weighting method is an essential component in the determination of the optimal size of a batch system in Friedman and Mehrez (1992) and a fuzzy optimal design problem concerning a bridge is solved in Ohkubo et al. (1998).

Linear problems with two objective functions are studied in Gass and Saaty (1955). The systematic generation of the Pareto optimal set is possible in these problems by parametric optimization. The Pareto optimal set of multiobjective optimization problems with convex quadratic objective functions and linear equality constraints is characterized analytically in Goh and Yang (1996). The characterization involves the weighting method and active set methods.

Systematic ways of perturbing the weights to obtain different Pareto optimal solutions are suggested in Chankong and Haimes (1983a, b, pp. 234–236). In addition, an algorithm for generating different weighting coefficients automatically for convex (nonlinear) problems to produce an approximation of the Pareto optimal set is proposed in Caballero et al. (1997).

A method for reducing the Pareto optimal set (of an MOLP problem) before it is presented to the decision maker is suggested in Soloveychik (1983). Pareto optimal solutions are first generated by the weighting method. Then, statistical analysis (factor analysis) is used to group and partition the Pareto optimal set into groups of relatively homogeneous elements. Finally, typical solutions from the groups are chosen and presented to the decision maker.

It is suggested in Koski and Silvennoinen (1987) that the weighting method can be used to reduce the number of the objective functions before the actual solution process. The original objective functions are divided into groups such that a linear combination of the objective functions in each group forms a new objective function, and these new objective functions form a new multiobjective optimization problem. It is stated that every Pareto optimal solution of the new problem is also a Pareto optimal solution of the original problem, but the reverse result is not generally true.

As mentioned earlier, the weighting vector that produces a certain Pareto optimal solution is not necessarily unique. This is particularly true for linear problems. A method is presented in Steuer (1986, pp. 183–187) for determining ranges for weighting vectors that produce the same solution. Note that some weighting vectors may produce unbounded single objective optimization problems. This does not mean that the problem may not have feasible solutions with some other weighting vectors.

A property related to producing different Pareto optimal solutions by altering the weighting coefficients is the weak stability of the system. On the one hand, a small change in the weighting coefficients may cause big changes in the objective vectors. On the other hand, dramatically different weighting coefficients may produce nearly similar objective vectors. The reason is that the weighting problem is not a Lipschitzian function of the weighting coefficients. In addition, it is emphasized and illustrated in Das and Dennis (1997) that an evenly distributed set of weighting vectors does not necessarily produce an evenly distributed representation of the Pareto optimal set, even if the problem is convex. Further, Das and Dennis demonstrate how an even spread of Pareto optimal solutions is obtained only for very special shapes of Pareto optimal sets. The treatment concerns two objective functions.

An entropy-based formulation of the weighting method is suggested in Sultan and Templeman (1996). The entropy-based objective function to be optimized has only one parameter no matter what the number of the original objective functions is. A representation of the Pareto optimal set can be generated by varying the value of the single parameter. The properties of the suggested method are the same as those of the weighting method, for example, all the Pareto optimal solutions of nonconvex problems cannot be found.

3.1.3. Weighting Method as an A Priori Method

The weighting method can be used so that the decision maker specifies a weighting vector representing her or his preference information. In this case, the weighting problem can be considered (a negative of) a value function (remember that value functions are maximized). Note that according to Remark 2.8.7 of Part I, the weighting coefficients provided by the decision maker are now nothing but marginal rates of substitution ($m_{ij} = w_j/w_i$). When the weighting method is used in this fashion, it can be considered to belong to the class of a priori methods. Related to this, a method for assisting in the determination of the weighting coefficients is presented in Batishchev et al. (1991). This method can also be extended into an interactive form by letting the decision maker modify the weighting vectors after each iteration.

The objective functions should be normalized or scaled so that their objective values are approximately of the same magnitude (see Subsection 2.4.3 in Part I). Only in this way can one control and manoeuvre the method to produce solutions of a desirable nature in proportion to the ranges of the ob-

jective functions. Otherwise the role of the weighting coefficients may be greatly misleading.

If the weighting method is used as an a priori method one can ask what the weighting coefficients in fact represent. Often they are said to reflect the relative importance of the objective functions. However, it is not at all clear what underlies this notion, as discussed in Roy and Mousseau (1996). It is remarked in Hobbs (1986) that instead of relative importance, the weighting coefficients should represent the rate at which the decision maker is willing to trade off values of the objective functions.

It must be noted that if some of the objective functions correlate with each other, then seemingly 'good' weighting vectors may produce poor results and seemingly 'bad' weighting vectors may produce useful results (see Steuer (1986, pp. 198–199) for an illustrative example).

On the basis of practical experience it is emphasized in Wierzbicki (1986b) that weighting coefficients are not easy to interpret and understand for average decision makers.

3.1.4. Concluding Remarks

The weighting method is a simple way to generate different Pareto optimal solutions. Pareto optimality is guaranteed if the weighting coefficients are positive or the solution is unique.

Applying Theorem 3.1.2, we know that different Pareto optimal solutions can be obtained by the weighting method by altering the positive weighting coefficients. However, in practical calculations the condition $w_i \geq \varepsilon$, where $\varepsilon > 0$, must be used instead of the condition $w_i > 0$ for all $i = 1, \ldots, k$. This necessitates a correct choice as to the value of ε. All the Pareto optimal solutions in some convex problems may be found if ε is small enough. But the concept of 'small enough' is problem-dependent and for this reason difficult to specify in advance, as pointed out in Korhonen and Wallenius (1989a).

As observed before, the weakness of the weighting method is that all of the Pareto optimal points cannot be found if the problem is nonconvex. If this is the case, a duality gap is said to occur (according to duality theory). The same weakness may also occur in problems with discontinuous objective functions as demonstrated in Kitagawa et al. (1982).

Sometimes, the results concerning the weighting method are presented in a simpler form, assuming that zeros are not accepted as weighting coefficients. It may seem that the weighting coefficient zero makes no sense. It means that we have included in the problem some objective function that has no significance at all. Nevertheless, zero values have here been included to make the presentation more general. On the other hand, by also allowing zeros as weighting coefficients, it is easy to explore how solutions change when some objective function is dropped.

Employing the weighting method as an a priori method presumes that the decision maker's underlying value function is or can be approximated by a linear

function (see Section 2.6 in Part I). This is in many cases a rather simplifying assumption. In addition, it must be noted that altering the weighting vectors linearly does not have to mean that the values of the objective functions also change linearly. It is, moreover, difficult to control the direction of the solutions by the weighting coefficients, as illustrated in Nakayama (1995).

3.2. ε-Constraint Method

In the ε-constraint method, introduced in Haimes et al. (1971), one of the objective functions is selected to be optimized and all the other objective functions are converted into constraints by setting an upper bound to each of them. The problem to be solved is now of the form

$$\begin{aligned} &\text{minimize} && f_\ell(\mathbf{x}) \\ &\text{subject to} && f_j(\mathbf{x}) \leq \varepsilon_j \quad \text{for all} \quad j = 1, \ldots, k, \ j \neq \ell, \\ & && \mathbf{x} \in S, \end{aligned} \tag{3.2.1}$$

where $\ell \in \{1, \ldots, k\}$. Problem (3.2.1) is called an *ε-constraint problem*.

An alternative formulation is proposed in Lin (1976a, b), where proper equality constraints are used instead of the above-mentioned inequalities. The solutions obtained by this proper equality method are Pareto optimal under certain assumptions. Here, however, we concentrate on formulation (3.2.1).

3.2.1. Theoretical Results on Weak and Pareto Optimality

We begin by proving a result concerning weak Pareto optimality.

Theorem 3.2.1. The solution of ε-constraint problem (3.2.1) is weakly Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a solution of the ε-constraint problem. Let us assume that $\mathbf{x}^*$ is not weakly Pareto optimal. In this case, there exists some other $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$.

This means that $f_j(\mathbf{x}) < f_j(\mathbf{x}^*) \leq \varepsilon_j$ for all $j = 1, \ldots, k$, $j \neq \ell$. Thus $\mathbf{x}$ is feasible with respect to the ε-constraint problem. While in addition $f_\ell(\mathbf{x}) < f_\ell(\mathbf{x}^*)$, we have a contradiction to the assumption that $\mathbf{x}^*$ is a solution of the ε-constraint problem. Thus, $\mathbf{x}^*$ has to be weakly Pareto optimal. □

Next, we handle Pareto optimality and the ε-constraint method.

Theorem 3.2.2. A decision vector $\mathbf{x}^* \in S$ is Pareto optimal if and only if it is a solution of ε-constraint problem (3.2.1) for every $\ell = 1, \ldots, k$, where $\varepsilon_j = f_j(\mathbf{x}^*)$ for $j = 1, \ldots, k$, $j \neq \ell$.

Proof. Necessity: Let $\mathbf{x}^* \in S$ be Pareto optimal. Let us assume that it does not solve the ε-constraint problem for some ℓ where $\varepsilon_j = f_j(\mathbf{x}^*)$ for $j = 1, \ldots, k$, $j \neq \ell$. Then there exists a solution $\mathbf{x} \in S$ such that $f_\ell(\mathbf{x}) < f_\ell(\mathbf{x}^*)$ and $f_j(\mathbf{x}) \leq f_j(\mathbf{x}^*)$ when $j \neq \ell$. This contradicts the Pareto optimality of $\mathbf{x}^*$. In other words, $\mathbf{x}^*$ has to solve the problem for any objective function.

Sufficiency: Since $\mathbf{x}^* \in S$ is by assumption a solution of the ε-constraint problem for every $\ell = 1, \ldots, k$, there is no $\mathbf{x} \in S$ such that $f_\ell(\mathbf{x}) < f_\ell(\mathbf{x}^*)$ and $f_j(\mathbf{x}) \leq f_j(\mathbf{x}^*)$ when $j \neq \ell$. This is the definition of Pareto optimality for $\mathbf{x}^*$. □

Note that according to the necessity component of Theorem 3.2.2, it is possible to find every Pareto optimal solution of any multiobjective optimization problem by the ε-constraint method (regardless of the convexity of the problem).

Theorem 3.2.3. A point $\mathbf{x}^* \in S$ is Pareto optimal if it is a unique solution of ε-constraint problem (3.2.1) for some ℓ with $\varepsilon_j = f_j(\mathbf{x}^*)$ for $j = 1, \ldots, k$, $j \neq \ell$.

Proof. Let $\mathbf{x}^* \in S$ be a unique solution of the ε-constraint problem for some ℓ. Let us assume that it is not Pareto optimal. In other words, there exists some point $\mathbf{x}^\circ \in S$ such that $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and for at least one index j is valid $f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*)$. The uniqueness of $\mathbf{x}^*$ means that for all $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$, $i \neq \ell$, is $f_\ell(\mathbf{x}^*) < f_\ell(\mathbf{x})$. Here we have a contradiction with the preceding inequalities, and $\mathbf{x}^*$ must be Pareto optimal. □

The following theorem is a straightforward extension of Theorem 3.2.3.

Theorem 3.2.4. The unique solution of ε-constraint problem (3.2.1) is Pareto optimal for any given upper bound vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{\ell-1}, \varepsilon_{\ell+1}, \ldots, \varepsilon_k)^T$.

Proof. Let $\mathbf{x}^* \in S$ be a unique solution of the ε-constraint problem. This means that $f_\ell(\mathbf{x}^*) < f_\ell(\mathbf{x})$ for all $\mathbf{x} \in S$ when $f_j(\mathbf{x}^*) \leq \varepsilon_j$ for every $j = 1, \ldots, k$, $j \neq \ell$. Let us assume that $\mathbf{x}^*$ is not Pareto optimal. In this case, there exists a vector $\mathbf{x}^\circ \in S$ such that $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and the inequality is strict for at least one index j.

If $j = \ell$, this means that $f_\ell(\mathbf{x}^\circ) < f_\ell(\mathbf{x}^*)$ and $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*) \leq \varepsilon_j$ for all $i \neq \ell$. Here we have a contradiction with the fact that $\mathbf{x}^*$ is a solution of the ε-constraint problem.

On the other hand, if $j \neq \ell$, then $f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*) \leq \varepsilon_j$, $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*) \leq \varepsilon_i$ for all $i \neq j$ and ℓ, and $f_\ell(\mathbf{x}^\circ) \leq f_\ell(\mathbf{x}^*)$. This is in contradiction to $\mathbf{x}^*$ as a unique solution of the ε-constraint problem, and $\mathbf{x}^*$ has to be Pareto optimal. □

In Figure 3.2.1, different upper bounds for the objective function f_2 are given, while the function f_1 is to be minimized. The Pareto optimal set is

shown by a bold line. The upper bound level ε^1 is too tight and so the feasible region is empty. On the other hand, the level ε^4 does not restrict the region at all. If it is used as the upper bound, the point $\mathbf{z}^4$ is obtained as a solution. It is Pareto optimal according to Theorem 3.2.4. Correspondingly, for the upper bound ε^3 the point $\mathbf{z}^3$ is obtained as a Pareto optimal solution. The point $\mathbf{z}^2$ is the optimal solution for the upper bound ε^2. Its Pareto optimality can be proved according to Theorem 3.2.3. Theorem 3.2.2 can be applied as well.

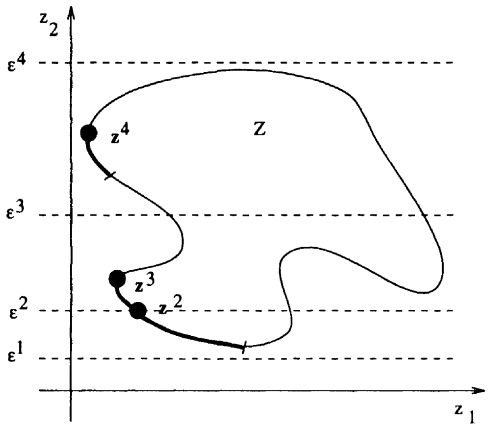

Figure 3.2.1. Different upper bounds for the ε-constraint method.

To ensure that a solution produced by the ε-constraint method is Pareto optimal, we have to either solve k different problems or obtain a unique solution. In general, uniqueness is not necessarily easy to verify. However, if for example, the problem is convex and the function f_ℓ to be minimized is strictly convex, we know that the solution is unique without further checking (see Chankong and Haimes (1983b, p. 131)).

According to Theorem 3.2.1 we know that the ε-constraint method produces weakly Pareto optimal solutions without any additional assumptions. We can show that any weakly Pareto optimal solution can be found with the ε-constraint method (for some objective function to be minimized) if the feasible region is convex and all the objective functions are quasiconvex and strictly quasiconvex. This result is derived in Ruíz-Canales and Rufián-Lizana (1995) and a shortened proof is also given in Luc and Schaible (1997). The value of this result is somewhat questionable because usually we are interested in Pareto optimal solutions and we know that any of them can be found with the ε-constraint method.

3.2.2. Connections with the Weighting Method

The relationships between the weighting method and the ε-constraint method are presented in the following theorems.

Theorem 3.2.5. Let $\mathbf{x}^* \in S$ be a solution of weighting problem (3.1.1) and $\mathbf{0} \leq \mathbf{w} \in \mathbf{R}^k$ be the corresponding weighting vector. Then

(1) if $w_\ell > 0$, $\mathbf{x}^*$ is a solution of the ε-constraint problem for f_ℓ as the objective function and $\varepsilon_j = f_j(\mathbf{x}^*)$ for $j = 1, \ldots, k$, $j \neq \ell$; or
(2) if $\mathbf{x}^*$ is a unique solution of weighting problem (3.1.1), then $\mathbf{x}^*$ is a solution of the ε-constraint problem when $\varepsilon_j = f_j(\mathbf{x}^*)$ for $j = 1, \ldots, k$, $j \neq \ell$ and for every f_ℓ, $\ell = 1, \ldots, k$, as the objective function.

Proof. Let $\mathbf{x}^* \in S$ be a solution of the weighting problem for some weighting vector $\mathbf{0} \leq \mathbf{w} \in \mathbf{R}^k$.

(1) We assume that $w_\ell > 0$. We have

$$\sum_{i=1}^{k} w_i f_i(\mathbf{x}) \geq \sum_{i=1}^{k} w_i f_i(\mathbf{x}^*) \tag{3.2.2}$$

for all $\mathbf{x} \in S$.

Let us suppose that $\mathbf{x}^*$ is not a solution of the ε-constraint problem. Then there exists a point $\hat{\mathbf{x}} \in S$ such that $f_\ell(\hat{\mathbf{x}}) < f_\ell(\mathbf{x}^*)$ and $f_j(\hat{\mathbf{x}}) \leq f_j(\mathbf{x}^*)$ when $j = 1, \ldots, k$, $j \neq \ell$. We assumed that $w_\ell > 0$ and $w_i \geq 0$ when $i \neq \ell$. Now we have

$$0 > w_\ell(f_\ell(\hat{\mathbf{x}}) - f_\ell(\mathbf{x}^*)) + \sum_{\substack{i=1 \\ i \neq \ell}}^{k} w_i(f_i(\hat{\mathbf{x}}) - f_i(\mathbf{x}^*)),$$

which is a contradiction with inequality (3.2.2). Thus $\mathbf{x}^*$ is a solution of the ε-constraint problem.

(2) If $\mathbf{x}^*$ is a unique solution of the weighting problem, then for all $\mathbf{x} \in S$

$$\sum_{i=1}^{k} w_i f_i(\mathbf{x}^*) < \sum_{i=1}^{k} w_i f_i(\mathbf{x}). \tag{3.2.3}$$

If there is some objective function f_ℓ such that $\mathbf{x}^*$ does not solve the ε-constraint problem when f_ℓ is to be minimized, then we can find a solution $\hat{\mathbf{x}} \in S$ such that $f_\ell(\hat{\mathbf{x}}) < f_\ell(\mathbf{x}^*)$ and $f_j(\hat{\mathbf{x}}) \leq f_j(\mathbf{x}^*)$ when $j \neq \ell$. This means that for any $\mathbf{w} \geq \mathbf{0}$ is $\sum_{i=1}^{k} w_i f_i(\hat{\mathbf{x}}) \leq \sum_{i=1}^{k} w_i f_i(\mathbf{x}^*)$. This contradicts inequality (3.2.3). Thus $\mathbf{x}^*$ is a solution of the ε-constraint problem for all f_ℓ to be minimized. □

The next theorem is valid for convex problems.

Theorem 3.2.6. Let the multiobjective optimization problem be convex. If $\mathbf{x}^* \in S$ is a solution of ε-constraint problem (3.2.1) for any given f_ℓ to be

minimized and $\varepsilon_j = f_j(\mathbf{x}^*)$ for $j = 1, \ldots, k$, $j \neq \ell$, then there exists a weighting vector $\mathbf{0} \leq \mathbf{w} \in \mathbf{R}^k$, $\sum_{i=1}^{k} w_i = 1$, such that $\mathbf{x}^*$ is also a solution of weighting problem (3.1.1).

Proof. The proof needs a so-called generalized Gordan theorem. See Chankong and Haimes (1983b, p. 121) and references therein.

We have now appropriate tools for proving Theorem 3.1.4 from the previous section.

Proof. (Proof of Theorem 3.1.4) Since $\mathbf{x}^*$ is Pareto optimal, it is by Theorem 3.2.2 a solution of the ε-constraint problem for every objective function f_ℓ to be minimized. The proof is completed with the aid of the convexity assumption and Theorem 3.2.6. □

A diagram representing several results concerning the characterization of Pareto optimal solutions and the optimality conditions of the weighting method, the ε-constraint method and a so-called jth Lagrangian method, their relations and connections is presented in Chankong and Haimes (1982, 1983b, pp. 119). The jth Lagrangian method, presented in Benson and Morin (1977), means solving the problem

$$\begin{aligned} &\text{minimize} \quad f_j(\mathbf{x}) + \sum_{\substack{i=1 \\ i \neq j}}^{k} u_i f_i(\mathbf{x}) \\ &\text{subject to} \quad \mathbf{x} \in S, \end{aligned} \tag{3.2.4}$$

where $\mathbf{u} = (u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_k)^T$ and $u_i \geq 0$ for all $i \neq j$. The jth Lagrangian method is from a computational viewpoint almost equal to the weighting method. This is why it is not studied more closely here. Chankong and Haimes have treated the problems separately to emphasize two ways of arriving at the same point.

3.2.3. Theoretical Results on Proper Pareto Optimality

Let us now return to the ε-constraint problem and the proper Pareto optimality of its solutions. In Benson and Morin (1977), an auxiliary function, called the *perturbation function*, $v\colon \mathbf{R}^{k-1} \to \mathbf{R}$ associated with the ε-constraint problem is defined in the form (modified here for the minimization problem)

$$v(\mathbf{y}) = \inf_{\mathbf{x} \in S} \{ f_\ell(\mathbf{x}) \mid f_j(\mathbf{x}) - \varepsilon_j \leq y_j \quad \text{for all} \quad j = 1, \ldots, k, \; j \neq \ell \}.$$

Naturally, the optimal value of the objective function of the ε-constraint problem is $v(\mathbf{0})$. We can now define the stability of ε-constraint problems.

Definition 3.2.7. The ε-constraint problem (3.2.1) is said to be *stable* when $v(\mathbf{0})$ is finite and there exists a scalar $R > 0$ such that, for all $\mathbf{0} \neq \mathbf{y} \in \mathbf{R}^{k-1}$

$$\frac{v(\mathbf{0}) - v(\mathbf{y})}{\|\mathbf{y}\|} \leq R.$$

After this, a theorem concerning the proper Pareto optimality of the solutions of the ε-constraint problem can be presented.

Theorem 3.2.8. Let the multiobjective optimization problem be convex and let $\mathbf{x}^* \in S$ be Pareto optimal. Then $\mathbf{x}^*$ is properly Pareto optimal if and only if ε-constraint problem (3.2.1) is stable for each $\ell = 1, \ldots, k$, where $\varepsilon_j = f_j(\mathbf{x}^*)$ for all $j = 1, \ldots, k$, $j \neq \ell$.

Proof. See Benson and Morin (1977) or Sawaragi et al. (1985, p. 88).

Let us now suppose that the feasible region is of the form

$$S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}.$$

The ε-constraint problem is a constrained single objective optimization problem and it can be converted into an unconstrained problem by formulating a *Lagrange function* of the form

$$f_\ell(\mathbf{x}) + \sum_{\substack{j=1 \\ j \neq \ell}}^{k} \lambda_j (f_j(\mathbf{x}) - \varepsilon_j) + \sum_{i=1}^{m} \mu_i g_i(\mathbf{x})$$

to be minimized. Setting some assumptions on the *Lagrange multipliers* $\boldsymbol{\lambda} \in \mathbf{R}^{k-1}$ and $\boldsymbol{\mu} \in \mathbf{R}^m$, we can derive more conditions for proper Pareto optimality.

In the following, we need the constraint qualification of Definition 3.1.10 of Part I, that is the definition of a regular point applied to the ε-constraint problem. In other words, a point $\mathbf{x}^* \in S$ is regular if the gradients of the active constraints of the ε-constraint problem at $\mathbf{x}^*$ are linearly independent.

For clarity, we shall now formulate the classical Karush-Kuhn-Tucker necessary condition for optimality (see Kuhn and Tucker (1951)) applied to the ε-constraint problem. The proof for general nonlinear problems is presented, for example, in Kuhn and Tucker (1951) and Luenberger (1984, p. 315). The condition can also be derived from the optimality conditions for multiobjective optimization problems, presented in Section 3.1 of Part I.

Note 3.2.9. (*Karush-Kuhn-Tucker necessary optimality condition applied to the ε-constraint problem*) Let the objective and the constraint functions be continuously differentiable at $\mathbf{x}^* \in S$ which is a regular point of the constraints of the ε-constraint problem. A necessary condition for $\mathbf{x}^*$ to be a solution of the

ε-constraint problem is that there exist vectors $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^{k-1}$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that

$$(1) \quad \nabla f_\ell(\mathbf{x}^*) + \sum_{\substack{j=1 \\ j \neq \ell}}^{k} \lambda_j \nabla (f_j(\mathbf{x}^*) - \varepsilon_j) + \sum_{i=1}^{m} \mu_i \nabla g_i(\mathbf{x}^*) = 0$$

$$(2) \quad \lambda_j (f_j(\mathbf{x}^*) - \varepsilon_j) = 0 \text{ for all } j \neq \ell, \quad \mu_i g_i(\mathbf{x}^*) = 0 \text{ for all } i = 1, \ldots, m.$$

Note that the (Lagrange) multipliers $\boldsymbol{\lambda}$ are in what follows called *Karush-Kuhn-Tucker multipliers*, when they are associated with the Karush-Kuhn-Tucker optimality condition. The condition states, for example, that if the constraint concerning f_j is not active, the corresponding multiplier λ_j must be equal to zero.

We can now present the following theorem.

Theorem 3.2.10. Let all the objective and the constraint functions be continuously differentiable at $\mathbf{x}^* \in S$ which is a regular point of the constraints of the ε-constraint problem. Then the following is valid.

(1) If $\mathbf{x}^*$ is properly Pareto optimal, then $\mathbf{x}^*$ solves the ε-constraint problem for some f_ℓ being minimized and $\varepsilon_j = f_j(\mathbf{x}^*)$ (for $j = 1, \ldots, k$, $j \neq \ell$) with all the Karush-Kuhn-Tucker multipliers associated with the constraints $f_j(\mathbf{x}) \leq \varepsilon_j$ for $j = 1, \ldots, k$, $j \neq \ell$, being positive.

(2) If the multiobjective optimization problem is convex, then $\mathbf{x}^*$ is properly Pareto optimal if it is a solution of the ε-constraint problem with the Karush-Kuhn-Tucker multipliers associated with the constraints $f_j(\mathbf{x}) \leq \varepsilon_j$ for $j = 1, \ldots, k$, $j \neq \ell$, being positive.

Proof. See Chankong and Haimes (1983b, p. 143).

It can also be proved that if some solution is improperly Pareto optimal and the problem is convex, then some of the associated Karush-Kuhn-Tucker multipliers equal zero. On the other hand, if some of the Karush-Kuhn-Tucker multipliers equal zero, then the solution of the ε-constraint problem is improperly Pareto optimal (see, e.g., Chankong and Haimes (1983a)).

According to Theorem 3.2.10, we can say that if a multiobjective optimization problem is solved by the ε-constraint method, proper Pareto optimality can be checked by employing the Lagrange function. In the previous section in connection with the weighting method, we also presented some conditions for proper Pareto optimality. Let us mention that proper Pareto optimality is characterized with the help of jth Lagrangian problem (3.2.4) in Benson and Morin (1977). There are, however, many methods where proper Pareto optimality is difficult to guarantee algorithmically.

3.2.4. Connections with Trade-Off Rates

The relationships between Karush-Kuhn-Tucker multipliers and trade-off rates are studied in Chankong and Haimes (1983b, pp. 159–165) and Haimes and Chankong (1979). Indeed, under certain conditions to be presented in the following, the Karush-Kuhn-Tucker multipliers of the Lagrange problem are equivalent to trade-off rates.

For notational convenience we state the second-order sufficient condition applied to the ε-constraint problem. See Chankong and Haimes (1983b, p. 58) for details.

Note 3.2.11. (*Second-order sufficient condition for optimality applied to the ε-constraint problem*) Let the objective and the constraint functions be twice continuously differentiable at $\mathbf{x}^* \in S$ which is a regular point of the constraints of the ε-constraint problem. A sufficient condition for $\mathbf{x}^*$ to be a solution of the ε-constraint problem is that there exist vectors $\mathbf{0} \leq \boldsymbol{\lambda} \in \mathbf{R}^{k-1}$ and $\mathbf{0} \leq \boldsymbol{\mu} \in \mathbf{R}^m$ such that the optimality condition of Note 3.2.9 is satisfied and the Hessian matrix of the corresponding Lagrange function

$$\nabla^2 f_\ell(\mathbf{x}^*) + \sum_{\substack{j=1 \\ j \neq \ell}}^{k} \lambda_j \nabla^2 (f_j(\mathbf{x}^*) - \varepsilon_j) + \sum_{i=1}^{m} \mu_i \nabla^2 g_i(\mathbf{x}^*)$$

is positive definite on the set $\{\mathbf{d} \in \mathbf{R}^n \mid \nabla g_i(\mathbf{x}^*)^T \mathbf{d} = 0 \text{ for all } i \text{ such that } \mu_i > 0\}$.

A connection between Karush-Kuhn-Tucker multipliers and trade-off rates is presented in the following theorem. The upper bound vector is denoted by $\boldsymbol{\varepsilon}^\circ \in \mathbf{R}^{k-1}$ and it is assumed to be chosen so that feasible solutions exist.

Theorem 3.2.12. Let $\mathbf{x}^* \in S$ be a solution of ε-constraint problem (3.2.1) for some f_ℓ, $\ell = 1, \ldots, k$, to be minimized. Let $\lambda_j = \lambda_{\ell j}$, for $j = 1, \ldots, k$, $j \neq \ell$, be the corresponding Karush-Kuhn-Tucker multipliers associated with the constraints $f_j(\mathbf{x}) \leq \varepsilon_j^\circ$ for $j \neq \ell$. If

(1) $\mathbf{x}^*$ is a regular point of the constraints of the ε-constraint problem,
(2) the second-order sufficient condition of Note 3.2.11 is satisfied at $\mathbf{x}^*$, and
(3) there are no degenerate constraints at $\mathbf{x}^*$ (i.e., the Karush-Kuhn-Tucker multipliers of all the constraints are strictly positive),

then $\lambda_{\ell j} = -\dfrac{\partial f_\ell(\mathbf{x}^*)}{\partial \varepsilon_j}$ for all $j = 1, \ldots, k$, $j \neq \ell$.

Proof. The proof is based on the implicit function theorem, see Luenberger (1984, p. 313).

From the assumption $\lambda_j(f_j(\mathbf{x}^*) - \varepsilon_j) = 0$ for all $j = 1, \ldots, k$, $j \neq \ell$ of the Karush-Kuhn-Tucker necessary optimality condition and the nondegeneracy of the constraints we know that $f_j(\mathbf{x}^*) = \varepsilon_j$ for all $j \neq \ell$. Thus, from Theorem 3.2.12 we have the trade-off rates

$$\lambda_{\ell j} = -\frac{\partial f_\ell(\mathbf{x}^*)}{\partial f_j} \quad \text{for all} \quad j \neq \ell.$$

An important result concerning the relationship between Karush-Kuhn-Tucker multipliers and trade-off rates in a more general situation, where zero-valued multipliers also are accepted, is presented in the following. For notational simplicity we now suppose that the function to be minimized in the ε-constraint problem is f_k (i.e., we set $f_\ell = f_k$). In addition, the upper bounds $\boldsymbol{\varepsilon}^\circ \in \mathbf{R}^{k-1}$ are assumed to be chosen so that feasible solutions exist. This does not lose any generality. For details and a more extensive form of the theorem we refer to Chankong and Haimes (1983b, pp. 161–163).

Let λ_{kj} be the Karush-Kuhn-Tucker multipliers associated with the constraints $f_j(\mathbf{x}) \leq \varepsilon_j^\circ$, $j = 1, \ldots, k-1$. Without loss of generality we can assume that the first p $(1 \leq p \leq k-1)$ of the multipliers are strictly positive (i.e., $\lambda_{kj} > 0$ for $j = 1, \ldots, p$) and the rest $k - 1 - p$ multipliers equal zero (i.e., $\lambda_{kj} = 0$ for $j = p+1, \ldots, k-1$). We denote the objective vector corresponding to $\mathbf{x}^*$ by $\mathbf{z}^* \in Z$.

Theorem 3.2.13. Let $\mathbf{x}^* \in S$ be a solution of ε-constraint problem (3.2.1) (when f_k is minimized) such that

(1) $\mathbf{x}^*$ is a regular point of the constraints of the ε-constraint problem,
(2) the second-order sufficient condition of Note 3.2.11 is satisfied at $\mathbf{x}^*$, and
(3) all the active constraints at $\mathbf{x}^*$ are nondegenerate.

Then we have the following.

1) If $p = k-1$, that is, all the multipliers λ_{kj} are strictly positive, then the Pareto optimal surface in the feasible objective region in the neighbourhood of $\mathbf{z}^*$ can be represented by a continuously differentiable function $\bar{f}_k$ such that for each $(z_1, z_2, \ldots, z_k)^T$ in the neighbourhood of $\mathbf{z}^*$ is $z_k = \bar{f}_k(z_1, z_2, \ldots, z_{k-1})$. Moreover, for all $1 \leq j \leq p = k-1$ is

$$\lambda_{kj} = -\frac{\partial \bar{f}_k}{\partial f_j}(z_1^*, z_2^*, \ldots, z_{k-1}^*).$$

Thus λ_{kj} represents the partial trade-off rate between f_k and f_j at $\mathbf{x}^*$.

2) If $1 \leq p < k-1$, that is, some of the multipliers λ_{kj} equal zero, then the Pareto optimal surface in the feasible objective region in the neighbourhood of $\mathbf{z}^*$ can be represented by continuously differentiable functions $z_j = \bar{f}_j(z_1, \ldots, z_p, \varepsilon_{p+1}^\circ, \ldots, \varepsilon_{k-1}^\circ)$ for $j = p+1, \ldots, k$. Moreover, for all $1 \leq i \leq p$ is

$$\lambda_{ki} = -\frac{\partial \bar{f}_k}{\partial f_i}(z_1^*, \ldots, z_p^*, \varepsilon_{p+1}^\circ, \ldots, \varepsilon_{k-1}^\circ) = \frac{\nabla f_k(\mathbf{x}^*)^T \mathbf{d}_i^*}{\nabla f_i(\mathbf{x}^*)^T \mathbf{d}_i^*},$$

where $\mathbf{d}_i^*$ is the direction of $\partial \mathbf{x}(\boldsymbol{\varepsilon}^\circ)/\partial \varepsilon_i$. In addition, for all $p+1 \leq j \leq k-1$ is

$$\frac{\partial \bar{f}_j}{\partial f_i}(z_1^*, \ldots, z_p^*, \varepsilon_{p+1}^\circ, \ldots, \varepsilon_{k-1}^\circ) = \nabla f_j(\mathbf{x}^*)^T \frac{\partial \mathbf{x}(\boldsymbol{\varepsilon}^\circ)}{\partial \varepsilon_i}.$$

Thus λ_{ki} represents the total trade-off rate between f_k and f_i at $\mathbf{x}^*$ in the direction of $\partial \mathbf{x}(\boldsymbol{\varepsilon}^\circ)/\partial \varepsilon_i$.

Proof. See Chankong and Haimes (1983b, pp. 163–165).

Let us now consider the contents of Theorem 3.2.13. Part 1) says that under the given conditions there are exactly $k-1$ degrees of freedom in specifying a point on the (locally) Pareto optimal surface in the objective space in the neighbourhood of $\mathbf{z}^*$. In other words, when the values for $z_1, z_2, \ldots, z_{k-1}$ have been chosen from the neighbourhood of $\mathbf{z}^*$, then the value for z_k can be calculated from the given function and the resulting point $\mathbf{z}$ will lie on the (locally) Pareto optimal surface in the objective space.

Part 2) of Theorem 3.2.13 extends the result of part 1) by relaxing the assumption that all the constraints $f_j(\mathbf{x}) \leq \varepsilon_j^\circ$, $j = 1, \ldots, k-1$, should be active and nondegenerate, that is, $\lambda_{kj} > 0$ for all $j = 1, \ldots, k-1$. When the number of nondegenerate constraints is p $(< k-1)$, then the degree of freedom in specifying a point on the (locally) Pareto optimal surface in the objective space in the neighbourhood of $\mathbf{z}^*$ is the number of nondegenerate active constraints (p). The results of Theorem 3.2.13 will be needed in Section 5.1 when the ε-constraint method is used as a part of an interactive method.

3.2.5. Applications and Extensions

Systematic ways of perturbing the upper bounds to obtain different Pareto optimal solutions are suggested in Chankong and Haimes (1983a, b, pp. 283–295). The ε-constraint method is used for generating Pareto optimal solutions in Osman and Ragab (1986b). Then the solutions are clustered and a global Pareto optimum is located.

Sensitivity analysis of the ε-constraint method is dealt with in Rarig and Haimes (1983). An index is defined approximating the standard deviation of the optimal solution. The objective and the constraint functions are not supposed to be known for a certainty.

Now that we have introduced two basic methods it is worthwhile to mention a method for nonlinear problems presented in Osman and Ragab (1986a). It combines features from both the weighting method and the ε-constraint method. The nonconvex feasible objective region is divided into convex and

nonconvex parts. The positive feature of the weighting method that the feasible region is not disturbed in the solution process is utilized in the convex parts, and the capability of the ε-constraint method to find all the Pareto optimal solutions is utilized in the nonconvex parts. Therefore, merits of both these basic methods are exploited.

A method related to the ε-constraint method is presented in Youness (1995). It generates the Pareto optimal set for problems with quasiconvex (and lower semicontinuous) objective functions. The method is based on level sets. If we consider an objective vector $\mathbf{z}^h$ and level sets $L_i(z_i^h) = \{\mathbf{x} \in S \mid f_i(\mathbf{x}) \leq z_i^h\}$ for $i = 1, \ldots, k$, and if we have $\cap_{i=1}^k L_i(z_i^h) = \{\mathbf{z}^h\}$, then the vector $\mathbf{z}^h$ is Pareto optimal.

An entropy-based formulation of the ε-constraint method is suggested in Sultan and Templeman (1996). The entropy-based objective function to be optimized has only one parameter no matter what the number of the original objective functions is. A representation of the Pareto optimal set can be generated by varying the value of the single parameter. The entropy-based function contains logarithms and exponential functions.

3.2.6. Concluding Remarks

Theoretically, every Pareto optimal solution of any multiobjective optimization problem can be found by the ε-constraint method by altering the upper bounds and the function to be minimized. It must be stressed that even duality gaps in nonconvex problems (see, e.g., Section 2.10 of Part I and Chankong and Haimes (1983b, pp. 135–136)) do not disturb the functioning of the ε-constraint method. However, computationally, the conditions set by Theorems 3.2.2, 3.2.3 and 3.2.4 are not always very practical. For example, according to Theorem 3.2.2, the ε-constraint problem needs to be solved k times for all f_ℓ as objective functions in order to generate one Pareto optimal solution. On the other hand, the uniqueness of the solution demanded in the other theorems is not always too easy to check either.

Computationally, the ε-constraint method is more laborious than the weighting method because the number of constraints increases. It may be difficult to specify appropriate upper bounds for the objective functions. The components of the ideal objective vector can be used to help in the specification. Then we can set $\varepsilon_j = z_j^* + e_j$ for $j = 1, \ldots, k$, $j \neq \ell$, where e_j is some relatively small positive real number that can be altered.

The ε-constraint method can also be used as an a priori method, where the decision maker specifies f_ℓ and the upper bounds. Then it can be characterized as an ad hoc method. It means that one can never be completely sure how to select the objective function and the upper bounds to obtain a desirable solution. This is a common weakness with the a priori weighting method.

3.3. Hybrid Method

At this point it is worth mentioning a method combining the weighting method and the ε-constraint method. This method is described in Corley (1980) and Wendell and Lee (1977) in slightly different forms. The name hybrid method is introduced in Chankong and Haimes (1983a, b).

The *hybrid problem* to be solved is

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{k} w_i f_i(\mathbf{x}) \\ &\text{subject to} && f_j(\mathbf{x}) \leq \varepsilon_j \quad \text{for all} \quad j = 1, \ldots, k, \\ & && \mathbf{x} \in S, \end{aligned} \tag{3.3.1}$$

where $w_i > 0$ for all $i = 1, \ldots, k$.

Notice that problem (3.3.1) is equivalent to problem (2.10.1) in Part I if we set $w_i = 1$ for every $i = 1, \ldots, k$. In Corley (1980), the problem is formulated in a more general setting with a pointed convex ordering cone defining efficiency. Optimality results were already handled in Section 2.10 of Part I. Nevertheless, we write them down here as well.

Theorem 3.3.1. The solution of hybrid problem (3.3.1) is Pareto optimal for any upper bound vector $\boldsymbol{\varepsilon} \in \mathbf{R}^k$. On the other hand, if $\mathbf{x}^* \in S$ is Pareto optimal, then it is a solution of problem (3.3.1) for $\boldsymbol{\varepsilon} = \mathbf{f}(\mathbf{x}^*)$.

Proof. See Corley (1980) or Wendell and Lee (1977).

The set of Pareto optimal solutions can be found by solving problem (3.3.1) with methods for parametric constraints (where the parameter is the vector of upper bounds $\boldsymbol{\varepsilon}$), see, for example, Rao (1984, pp. 418–421). This means that the weighting coefficients do not have to be altered.

Optimality conditions for the solutions of problem (3.3.1) to be properly Pareto optimal are presented in Wendell and Lee (1977).

We can say that the positive features of the weighting method and the ε-constraint method are combined in the hybrid method. Namely, any Pareto optimal solution can be found independently of the convexity of the problem and one does not have to solve several problems or think about uniqueness to guarantee the Pareto optimality of the solutions. On the other hand, the specification of the parameter values may still be difficult. Computationally, the hybrid method is similar to the ε-constraint method (with the increased number of constraint functions).

3.4. Method of Weighted Metrics

In the method of the global criterion, introduced in Section 2.1, L_p- and L_∞-metrics were used to generate (weakly) Pareto optimal solutions. These metrics can also be weighted in order to produce different (weakly) Pareto optimal solutions. The weighted approach is also sometimes called compromise programming (see Zeleny (1973)). Here we use the term the method of weighted metrics.

3.4.1. Introduction

We assume that $w_i \geq 0$ for all $i = 1, \ldots, k$ and $\sum_{i=1}^{k} w_i = 1$. We obtain different solutions by altering the weighting coefficients w_i in the weighted L_p- and Tchebycheff metrics. The *weighted L_p-problem* for minimizing distances is now of the form

$$\begin{aligned} &\text{minimize} && \left(\sum_{i=1}^{k} w_i \left| f_i(\mathbf{x}) - z_i^\star \right|^p \right)^{1/p} \\ &\text{subject to} && \mathbf{x} \in S \end{aligned} \tag{3.4.1}$$

for $1 \leq p < \infty$. The *weighted Tchebycheff problem* is of the form

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} \left[w_i \left| f_i(\mathbf{x}) - z_i^\star \right| \right] \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{3.4.2}$$

Problem (3.4.2) was originally introduced in Bowman (1976). Again, denominators may be included. Further, the absolute value signs can be dropped because of the definition of the ideal objective vector, if it is known globally. Weighting vectors can also be used in connection with problems of form (2.1.4).

If $p = 1$, the sum of weighted deviations is minimized and the problem to be solved is equal to the weighting problem except for a constant (if $\mathbf{z}^\star$ is known globally). If $p = 2$, we have a method of least squares. When p gets larger, the minimization of the largest deviation becomes more and more important. Finally, when $p = \infty$, the only thing that matters is the weighted relative deviation of one objective function.

Problem (3.4.2) is nondifferentiable like its unweighted counterpart. Correspondingly, it can be solved in a differentiable form as long as the objective and the constraint functions are differentiable and $\mathbf{z}^\star$ is known globally. In this case, instead of problem (3.4.2), the problem

$$\begin{aligned} &\text{minimize} && \alpha \\ &\text{subject to} && \alpha \geq w_i \left(f_i(\mathbf{x}) - z_i^\star \right) \quad \text{for all} \quad i = 1, \ldots, k, \\ & && \mathbf{x} \in S, \end{aligned} \tag{3.4.3}$$

is solved, where both $\mathbf{x} \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$ are variables. This formulation will be utilized later.

3.4.2. Theoretical Results

In the following, we present some results concerning the weighted metrics. Most of the proofs are so closely reminiscent of those presented earlier in Section 2.1 that there is no reason to repeat them. We assume that the global ideal objective vector is known.

Theorem 3.4.1. The solution of weighted L_p-problem (3.4.1) (when $1 \leq p < \infty$) is Pareto optimal if either

(i) the solution is unique or
(ii) all the weighting coefficients are positive.

Proof. The proof is not presented here since it follows directly from the proofs of Theorems 3.1.2, 3.1.3 and 2.1.1. See Chankong and Haimes (1983b, p. 144) or Yu (1973).

Theorem 3.4.2. The solution of weighted Tchebycheff problem (3.4.2) is weakly Pareto optimal if all the weighting coefficients are positive.

Proof. The proof is a straightforward modification of the proof of Theorem 2.1.2.

Theorem 3.4.3. Weighted Tchebycheff problem (3.4.2) has at least one Pareto optimal solution.

Proof. The proof follows directly from the proof of Theorem 2.1.3.

Corollary 3.4.4. If weighted Tchebycheff problem (3.4.2) has a unique solution, it is Pareto optimal.

Convexity of the multiobjective optimization problem is needed in order to guarantee that every Pareto optimal solution can be found by the weighted L_p-problem (see Sawaragi et al. (1985, p. 81)). The following theorem shows that, on the other hand, every Pareto optimal solution can be found by the weighted Tchebycheff problem.

Theorem 3.4.5. Let $\mathbf{x}^* \in S$ be Pareto optimal. Then there exists a weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ such that $\mathbf{x}^*$ is a solution of weighted Tchebycheff problem (3.4.2), where the reference point is the utopian objective vector $\mathbf{z}^{**}$.

Proof. Let $\mathbf{x}^* \in S$ be Pareto optimal. Let us assume that there does not exist a weighting vector $\mathbf{w} > \mathbf{0}$ such that $\mathbf{x}^*$ is a solution of the weighted Tchebycheff problem. We know that $f_i(\mathbf{x}) > z_i^{**}$ for all $i = 1, \ldots, k$ and for all $\mathbf{x} \in S$. Now we choose $w_i = \beta/(f_i(\mathbf{x}^*) - z_i^{**})$ for all $i = 1, \ldots, k$, where $\beta > 0$ is some normalizing factor.

If $\mathbf{x}^*$ is not a solution of the weighted Tchebycheff problem, there exists another point $\mathbf{x}^\circ \in S$ that is a solution of the weighted Tchebycheff problem, meaning that

$$\max_i [w_i(f_i(\mathbf{x}^\circ) - z_i^{**})] < \max_i [w_i(f_i(\mathbf{x}^*) - z_i^{**})]$$
$$= \max_i \left[\frac{\beta}{f_i(\mathbf{x}^*) - z_i^{**}} (f_i(\mathbf{x}^*) - z_i^{**}) \right] = \beta.$$

Thus $w_i(f_i(\mathbf{x}^\circ) - z_i^{**}) < \beta$ for all $i = 1, \ldots, k$. This means that

$$\frac{\beta}{f_i(\mathbf{x}^*) - z_i^{**}} (f_i(\mathbf{x}^\circ) - z_i^{**}) < \beta,$$

and after simplifying the expression we have

$$f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$$

for all $i = 1, \ldots, k$. Here we have a contradiction with the Pareto optimality of $\mathbf{x}^*$, which completes the proof. □

A theorem, parallel to Theorem 3.4.5, is proved in Kaliszewski (1995).

3.4.3. Comments

Theorem 3.4.5 above sounds quite promising for the weighted Tchebycheff problem. Unfortunately, this is not the whole truth. In addition to the fact that every Pareto optimal solution can be found, weakly Pareto optimal solutions may also be included. Auxiliary calculation is needed in order to identify the weak ones. Remember that as far as the weighted L_p-problem ($1 \leq p < \infty$) is concerned, it produces Pareto optimal solutions but does not necessarily find all of them.

Selecting the value for the exponent p is treated in Ballestero (1997b) from the point of view of risk aversion. The conclusion is that for greater risk aversion we should use greater values for p. Another guideline is that for a smaller number of objective functions we should select greater p values.

More results concerning the properties of the L_p-metrics ($1 \leq p \leq \infty$) with and without the weighting coefficients can be found, for example, in Bowman (1976), Chankong and Haimes (1983b, pp. 144–146), Koski and Silvennoinen (1987), Nakayama (1985a) and Yu (1973), the first of these treating especially the Tchebycheff metric. Some results concerning the proper efficiency (in the sense of Henig) of the solutions of the weighted L_p-problem are presented briefly in Wierzbicki (1986b).

3.4.4. Connections with Trade-Off Rates

Useful results concerning trade-off rates and the weighted Tchebycheff problem are proved in Yano and Sakawa (1987). The approach is closely related to what was presented in Subsection 3.2.4 in connection with the ε-constraint problem.

Let us once again suppose that the feasible region is of the form

$$S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}.$$

All the objective and the constraint functions are assumed to be twice continuously differentiable, which is why problem (3.4.3) is the one to be dealt with.

Problem (3.4.3) is first formulated as an unconstrained minimization problem with one objective function, the Lagrange function, of the form

$$\alpha + \sum_{i=1}^{k} \lambda_i \big(w_i(f_i(\mathbf{x}) - z_i^\star) - \alpha\big) + \sum_{i=1}^{m} \mu_i g_i(\mathbf{x}), \tag{3.4.4}$$

where $\boldsymbol{\lambda} \in \mathbf{R}^k$ and $\boldsymbol{\mu} \in \mathbf{R}^m$ are Karush-Kuhn-Tucker multipliers. The decision variable vector being $(\alpha, \mathbf{x}) \in \mathbf{R}^{n+1}$, let us denote the minimal solution of function (3.4.4) by $\mathbf{y}^* \in \mathbf{R}^{n+1}$.

It is assumed that the assumptions in Theorem 3.2.13 when applied to problem (3.4.3) are satisfied. This means that $\mathbf{y}^*$ is a regular point of the constraints of (3.4.3), the second-order sufficient condition of Note 3.2.11 applied to problem (3.4.3) is satisfied at $\mathbf{y}^*$ and all the active constraints at $\mathbf{y}^*$ are nondegenerate. (The last assumption means that the Karush-Kuhn-Tucker multipliers of all the active constraints are positive.)

If all the constraints connected to the objective functions are active, we have

$$\lambda_{ij} = \frac{\lambda_j w_j}{\lambda_i w_i}.$$

Notice that the weighting coefficients have an essential role in these trade-off rates, unlike those related to the ε-constraint method. The procedure is not treated here in more detail because of its resemblance to what was presented in Subsection 3.2.4. For details, see Yano and Sakawa (1987).

3.4.5. Variants of the Weighted Tchebycheff Problem

Thus far, it has been proved that the weighted Tchebycheff problem can find any Pareto optimal solution. According to Corollary 3.4.4, the unique solution of the weighted Tchebycheff problem is Pareto optimal. If the solution is not unique or the uniqueness is difficult to guarantee, the weakness of the problem is that it may produce weakly Pareto optimal solutions as well. This weakness can be overcome in different ways. One possibility is to solve some additional

problem, for example problem (2.10.1) or problem (2.10.2), given in Part I. Problem (2.10.1) can also be modified so that instead of the sum of the objective functions, the sum of objective functions minus utopian objective values is minimized. Such an approach is handled in Section 5.4. It is also discussed, for example, in Korhonen (1997). One more way is to use lexicographic ordering (to be introduced in Section 4.2).

If additional optimizations must be avoided, another possibility is to vary the metric. Weakly Pareto optimal solutions can be avoided by giving a slight slope to the contour of the metric. The price to be paid is that in some cases it may be impossible to find every Pareto optimal solution. For that reason, properly Pareto optimal solutions are of interest here. Note that the utopian objective vector is used as a reference point as in Theorem 3.4.5.

It is suggested in Steuer (1986) and Steuer and Choo (1983) that the weighted Tchebycheff problem be varied by an augmentation term. In this case, the distance between the utopian objective vector and the feasible objective region is measured by an *augmented weighted Tchebycheff metric*. The *augmented weighted Tchebycheff problem* is of the form

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} \left[w_i |f_i(\mathbf{x}) - z_i^{\star\star}| \right] + \rho \sum_{i=1}^{k} |f_i(\mathbf{x}) - z_i^{\star\star}| \\ &\text{subject to} && \mathbf{x} \in S, \end{aligned} \tag{3.4.5}$$

where ρ is a sufficiently small positive scalar.

A slightly different *modified weighted Tchebycheff metric* is used in the *modified weighted Tchebycheff problem*

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} \left[w_i \left(|f_i(\mathbf{x}) - z_i^{\star\star}| + \rho \sum_{i=1}^{k} |f_i(\mathbf{x}) - z_i^{\star\star}| \right) \right] \\ &\text{subject to} && \mathbf{x} \in S, \end{aligned} \tag{3.4.6}$$

where ρ is a sufficiently small positive scalar. It is shown in Kaliszewski (1987) that problem (3.4.6) is equivalent (up to scalar multiplication) to that proposed in Choo and Atkins (1983).

The difference between the augmented and the modified weighted Tchebycheff problems is in the way the slope takes place in the metrics, as illustrated in Figure 3.4.1. In the augmented weighted Tchebycheff problem the slope is a function of the weighting coefficients and the parameter ρ. In other words, the slope may be different for each objective function, that is, for each coordinate of the objective space. As far as the modified weighted Tchebycheff problem is concerned, the slope is a function of the parameter ρ and, thus, constant for all the objective functions. In Figure 3.4.1 we have

$$\beta_i = \arctan \frac{\rho}{1 - w_i + \rho} \quad \text{and} \quad \beta = \arctan \frac{\rho}{1 + \rho}.$$

To ease the comparison, the dotted lines represent the slope of the weighted Tchebycheff metric. See Kaliszewski (1986, 1987) for details.

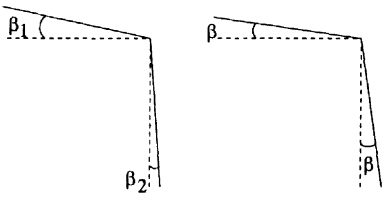

Figure 3.4.1. Slopes of two metrics.

The augmented weighted Tchebycheff problem is illustrated in Figure 3.4.2, where the dotted lines represent the contours of the augmented metric. The contour of the weighted Tchebycheff metric (continuous line) has only been added to ease the comparison. Sensitivity analysis of the augmented and the modified weighted Tchebycheff metrics is handled in Kaliszewski (1994, pp. 113–119).

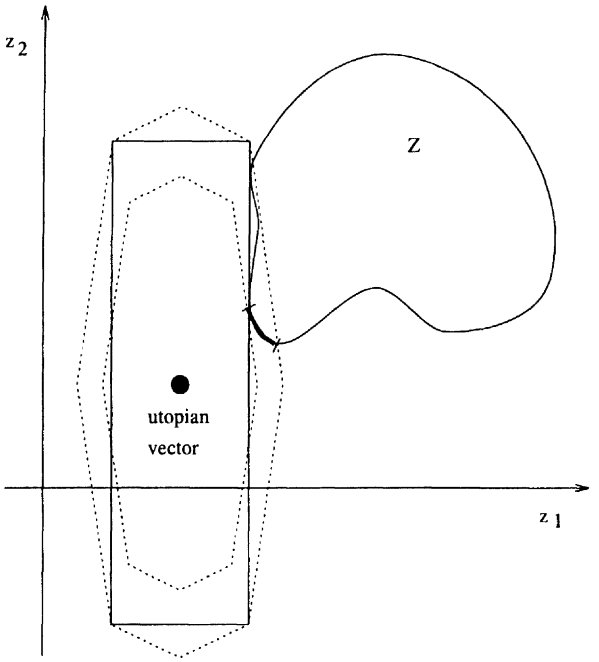

Figure 3.4.2. Augmented weighted Tchebycheff problem.

It is valid for both the augmented and the modified weighted Tchebycheff problem that they generate only properly Pareto optimal solutions and any properly Pareto optimal solution can be found. In what follows, the symbol M is the scalar from Definition 2.9.1 of proper Pareto optimality in Part I.

Theorem 3.4.6. A decision vector $\mathbf{x}^* \in S$ is properly Pareto optimal if and only if there exist a weighting vector $\mathbf{w} \in \mathbf{R}^k$ with $w_i > 0$ for all $i = 1, \ldots, k$ and a number $\rho > 0$ such that $\mathbf{x}^*$ is a solution of augmented weighted Tchebycheff problem (3.4.5).

In addition, for each properly Pareto optimal solution $\mathbf{x}^* \in S$ there exists a weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ such that $\mathbf{x}^*$ is a unique solution of problem (3.4.5) for every $\rho > 0$ satisfying $M \leq \min_{i=1,\ldots,k} w_i/((k-1)\rho)$.

Further, the inequality

$$M \leq \frac{1}{\rho} \max_{i=1,\ldots,k} w_i + (k-1)\rho$$

is valid for every solution $\mathbf{x}^*$ of problem (3.4.5).

Proof. See Kaliszewski (1994, pp. 51–53).

The proof in Kaliszewski (1994) is based on a cone separation technique. The necessity and the sufficiency components are also proved in Kaliszewski (1985).

The theorem for the modified weighted Tchebycheff problem is almost similar.

Theorem 3.4.7. A decision vector $\mathbf{x}^* \in S$ is properly Pareto optimal if and only if there exist a weighting vector $\mathbf{w} \in \mathbf{R}^k$ with $w_i > 0$ for all $i = 1, \ldots, k$ and a number $\rho > 0$ such that $\mathbf{x}^*$ is an optimal solution of modified weighted Tchebycheff problem (3.4.6).

In addition, for each properly Pareto optimal solution $\mathbf{x}^* \in S$ there exists a weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ such that $\mathbf{x}^*$ is a unique solution of problem (3.4.6) for every $\rho > 0$ satisfying $M \leq 1/((k-1)\rho)$.

Further, the inequality $M \leq (1 + (k-1)\rho)/\rho$ is valid for every solution $\mathbf{x}^*$ of problem (3.4.6).

Proof. See Kaliszewski (1994, pp. 48–50).

This necessity and sufficiency formulation is an extension of the original theorem in Choo and Atkins (1983). The necessary conditions in Theorems 3.4.6 and 3.4.7 are also proved in Kaliszewski (1995).

3.4.6. Connections with Global Trade-Offs

Some metrics for measuring the distance between the utopian objective vector and the feasible objective region can be formed in such a way that they produce solutions with selectively bounded global trade-offs. This is in reverse to the general way where trade-offs are calculated only after solutions have been generated.

For the simplicity of notations we here assume the global ideal objective vector and, thus, the global utopian objective vector to be known. This implies that we can drop the absolute value signs.

All the properly Pareto optimal solutions produced with modified weighted Tchebycheff problem (3.4.6) have bounded global trade-offs. Further, we have a common bound for every global trade-off involved.

Theorem 3.4.8. Let $\mathbf{x}^* \in S$ be a solution of modified weighted Tchebycheff problem (3.4.6) for some weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ and $\rho > 0$. In this case, the global trade-offs are bounded, that is

$$\Lambda_{ij}^G(\mathbf{x}^*) \leq \frac{1+\rho}{\rho}$$

for every $i,\ j = 1, \ldots, k,\ i \neq j$.

Proof. See Kaliszewski (1994, pp. 94–95).

Corresponding results can be proved for other types of problems, see Kaliszewski (1994, pp. 82–113).

Sometimes the decision maker may wish to set a priori bounds on some specific global trade-offs. Such a request calls for a scalarizing function of a special form. These topics are treated in Kaliszewski and Michalowski (1995, 1997). Thus far, the additional term multiplied with ρ was added to guarantee the proper Pareto optimality of the solutions. If we leave it out, we obtain weighted Tchebycheff problem (3.4.2) and, thus, weakly Pareto optimal solutions. In what follows, we use metrics without modification or augmentation terms but use other parameters σ and $\sigma_i > 0$ to control the bounds of the global trade-offs involved. Thus, the following results deal with weak Pareto optimality.

The next theorem handles a case where we wish to set a priori bounds for a group of selected global trade-offs. Let us choose a subset of the objective functions $I_0 \subset I = \{1, \ldots, k\}$ and define $I(i) = \{j \mid j \in I_0,\ j \neq i\}$.

Theorem 3.4.9. A decision vector $\mathbf{x}^* \in S$ is weakly Pareto optimal and

$$\Lambda_{ji}^G(\mathbf{x}^*) \leq \frac{1+\sigma}{\sigma} \quad \text{for all} \quad i \in I_0 \quad \text{and all} \quad j \in I(i) \tag{3.4.7}$$

if and only if there exist a weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ and a number $\sigma > 0$ such that $\mathbf{x}^*$ is a solution of the problem

$$\text{minimize} \quad \max\left[\max_{i \in I_0}\left[w_i\Big((1+\sigma)(f_i(\mathbf{x}) - z_i^{**}) + \sigma \sum_{j \in I(i)} \big(f_j(\mathbf{x}) - z_j^{**}\big)\Big)\right],\right.$$

$$\left.\max_{i \in I \setminus I_0}\left[w_i(f_i(\mathbf{x}) - z_i^{**})\right]\right]$$

$$\text{subject to} \quad \mathbf{x} \in S.$$

Proof. See Kaliszewski and Michalowski (1995).

Result (3.4.7) of Theorem 3.4.9 is utilized so that the decision maker is asked to specify upper bounds for the selected global trade-offs. These values are set as upper bounds to $(1+\sigma)/\sigma$. A lower bound for the parameter σ is obtained from these inequalities. By using the calculated σ value, we receive different weakly Pareto optimal solutions satisfying the global trade-off bounds by altering the weighting coefficients. In other words, we avoid generating solutions exceeding the specified bounds for global trade-off.

In Theorem 3.4.9 we have a common bound for the selected set of global trade-offs. This can further be generalized by using several different parameters σ.

Theorem 3.4.10. A decision vector $\mathbf{x}^* \in S$ is weakly Pareto optimal and for each $i \in I_0$ and each $j \in I_0$, $\Lambda_{ji}^G(\mathbf{x}^*)$ is bounded from above by a positive finite number if and only if there exist a weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ and numbers $\sigma_j > 0$ for $j \in I_0$, such that $\mathbf{x}^*$ is a solution of the problem

$$\begin{aligned} \text{minimize} \quad & \max\left[\max_{i\in I_0}\left[w_i\Big(f_i(\mathbf{x}) - z_i^{**} + \sum_{j\in I_0}\sigma_j\big(f_j(\mathbf{x}) - z_j^{**}\big)\Big)\right],\right. \\ & \left.\max_{i\in I\setminus I_0}\left[w_i(f_i(\mathbf{x}) - z_i^{**})\right]\right] \\ \text{subject to} \quad & \mathbf{x} \in S. \end{aligned} \tag{3.4.8}$$

Further, for any solution $\mathbf{x}^*$ of problem (3.4.8), we have

$$\Lambda_{ji}^G(\mathbf{x}^*) \leq \frac{1+\sigma_i}{\sigma_j}$$

for $i,\ j \in I_0$ and $j \neq i$.

Proof. See Kaliszewski and Michalowski (1997).

Theorem 3.4.10 is applied in the following way. If we want to generate weakly Pareto optimal solutions such that certain global trade-offs are bounded (each global trade-off with an individual bound), we form a system of equations from the global trade-off information. That is, we set $(1+\sigma_i)/\sigma_j$ equal to the specified upper bound, where desired. If the system is consistent, we solve it and obtain values for the parameters σ_j. If the system is inconsistent, some equation(s) must be dropped in order to form a consistent system. In this way, parameters σ_j are used to control the bounds of the selected global trade-offs.

Let us return to proper Pareto optimality. Theorem 3.4.8 can be modified to handle individual global trade-offs. This means that we solve a problem

$$\text{(3.4.9)} \qquad \begin{aligned} &\text{minimize} \quad \max_{i=1,\ldots,k} \left[w_i \Big(f_i(\mathbf{x}) - z_i^{\star\star} + \sum_{i=1}^{k} \rho_i (f_i(\mathbf{x}) - z_i^{\star\star}) \Big) \right] \\ &\text{subject to} \quad \mathbf{x} \in S. \end{aligned}$$

It is proved in Kaliszewski and Michalowski (1997) that Theorem 3.4.7 is valid with $\rho_i > 0$, $i = 1, \ldots, k$. In other words, a decision vector $\mathbf{x}^* \in S$ is properly Pareto optimal if and only if there exist a weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ and numbers $\rho_i > 0$ for all $i = 1, \ldots, k$ such that $\mathbf{x}^*$ is a solution of problem (3.4.9). Further, for the solution $\mathbf{x}^*$ of problem (3.4.9) the upper bounds of the global trade-offs are of the form

$$\Lambda_{ji}^{G}(\mathbf{x}^*) \leq \frac{1 + \rho_i}{\rho_j}$$

for $i,\ j = 1, \ldots, k$ and $i \neq j$.

3.4.7. Applications and Extensions

A shape optimization problem of a spillway profile is solved by the weighted L_2-metric (with denominators) in Wang and Zhou (1990). Further, an extension of the method of weighted metrics called composite programming is presented in Bárdossy et al. (1985). The L_p-metric ($p < \infty$) is divided into nested parts with different exponents. In particular, this approach can be applied to problems where objective functions consist of several components. The method is applied to problems of multiobjective watershed management and observation network design.

The weighted Tchebycheff metric is used in Kostreva et al. (1995) in deriving an integral approach for solving nonlinear problems with discontinuous objective functions and a discontinuous feasible region. The close connections between the weighted Tchebycheff metric and the weighting method for convex problems are handled in Dauer and Osman (1985). Karush-Kuhn-Tucker optimality conditions for these two methods are treated as well.

One more possibility of avoiding weakly Pareto optimal solutions is suggested in Helbig (1991). Helbig defines optimality with ordering cones. The idea is to perturb the ordering cone so that even though the method produces weakly efficient solutions, they are efficient to the original problem.

One way to apply the weighted Tchebycheff metric and its augmented version successfully will be introduced in Section 5.4.

3.4.8. Concluding Remarks

Particularly the method of weighted Tchebycheff metric and its variants are popular methods for generating Pareto optimal solutions. They work for convex as well as nonconvex problems (unlike the weighting method) and alteration of the parameters is easier than in the ε-constraint method.

As far as the weighted L_p-metrics are concerned, the solutions produced are Pareto optimal but not necessarily all of them are found (depending on the degree of nonconvexity of the problem). The weighted Tchebycheff metric can generate any Pareto optimal solution to any type of a problem. The drawback of also generating weakly Pareto optimal solutions can be overcome by augmenting or modifying the metric or by solving another optimization problem after minimizing the distance. These alternatives will be further addressed in Section 5.4. For nonconvex problems the success of the method of weighted metrics depends on whether the global ideal objective vector is known or not.

An interesting feature related to some variants of the weighted Tchebycheff metric is the ability to produce solutions with a priori-specified bounds for global trade-offs. In this way, a subset of weakly or properly Pareto optimal solutions can be generated satisfying given fixed bounds.

3.5. Achievement Scalarizing Function Approach

The approach to be presented is related to that of weighted metrics. Namely, we handle special types of scalarizing functions, termed achievement scalarizing functions. They have been introduced by Wierzbicki, for example, in Wierzbicki (1981, 1982, 1986a, b) (and are also handled in Wierzbicki (1977, 1980a, b)). Somewhat similar results for scalarizing functions are also presented in Jahn (1984) and Luc (1986).

3.5.1. Introduction

In the method of weighted L_p-metric or the weighted Tchebycheff metric, the distance is minimized between the ideal objective vector and the feasible objective region. If the global ideal objective vector is unknown, we may fail in producing (weakly) Pareto optimal solutions. In other words, if the reference point used is an objective vector inside the feasible objective region, the minimal distance to it is zero and we do not obtain a (weakly) Pareto optimal solution. We can overcome this weakness by replacing metrics with achievement scalarizing functions.

For example, weakly Pareto optimal solutions can be generated with any reference point $\bar{\mathbf{z}} \in \mathbf{R}^k$ by solving the problem

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} [\, w_i(f_i(\mathbf{x}) - \bar{z}_i)\,] \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{3.5.1}$$

It differs from weighted Tchebycheff problem (3.4.2) only in that the absolute value signs are missing. This change ensures that weakly Pareto optimal solutions are produced independently of the feasibility or infeasibility of the reference point.

Scalarizing functions of a special type are called achievement scalarizing functions. Problem (3.5.1) is one example of them. Let us next handle achievement scalarizing functions in general.

It is shown, for instance, in Wierzbicki (1980b, 1986a, b) that Pareto optimal solutions can be characterized by achievement scalarizing functions if the functions satisfy certain requirements. An *achievement scalarizing function* is a function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$, where $\bar{\mathbf{z}} \in \mathbf{R}^k$ is an arbitrary reference point of aspiration levels (see Definition 2.3.1 in Part I). In the following, we shorten the name to an *achievement function*.

Because we do not know the feasible objective region Z explicitly, in practice we minimize the function $s_{\bar{\mathbf{z}}}(\mathbf{f}(\mathbf{x}))$ subject to $\mathbf{x} \in S$ (see, e.g., Figure 2.2.1 in Part I). Thus, we deal with the feasible region in the decision variable space. For notational convenience, we, however, present the problem here as if it were solved in the feasible objective region.

3.5.2. Theoretical Results

We need to apply some of the general properties introduced in Part I to an achievement function $s_{\bar{\mathbf{z}}}$, namely the definitions of strictly increasing (Definition 2.1.8), strongly increasing (Definition 2.1.9) and ε-strongly increasing (Definition 2.1.10) functions. In the last-mentioned concept the definition of the set $\mathbf{R}^k_\varepsilon$ is the same as in connection with ε-proper Pareto optimality (see Definition 2.9.2 in Part I).

Next we can define order-representing and order-approximating achievement functions.

Definition 3.5.1. A continuous achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is *order-representing* if it is strictly increasing as a function of $\mathbf{z} \in Z$ for any $\bar{\mathbf{z}} \in \mathbf{R}^k$ and if

$$\{\mathbf{z} \in \mathbf{R}^k \mid s_{\bar{\mathbf{z}}}(\mathbf{z}) < 0\} = \bar{\mathbf{z}} - \operatorname{int} \mathbf{R}^k_+$$

(for all $\bar{\mathbf{z}} \in \mathbf{R}^k$).

Definition 3.5.2. A continuous achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is *order-approximating* if it is strongly increasing as a function of $\mathbf{z} \in Z$ for any $\bar{\mathbf{z}} \in \mathbf{R}^k$ and if

$$\bar{\mathbf{z}} - \mathbf{R}^k_{\bar{\varepsilon}} \subset \{\mathbf{z} \in \mathbf{R}^k \mid s_{\bar{\mathbf{z}}}(\mathbf{z}) \leq 0\} \subset \bar{\mathbf{z}} - \mathbf{R}^k_\varepsilon$$

(for all $\bar{\mathbf{z}} \in \mathbf{R}^k$) with $\varepsilon > \bar{\varepsilon} \geq 0$.

Remark 3.5.3. For a continuous order-representing or order-approximating achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ we have

$$s_{\bar{\mathbf{z}}}(\bar{\mathbf{z}}) = 0.$$

We can now present some optimality results concerning different types of achievement functions according to Wierzbicki (1986a, b). The *achievement problem* to be solved is

$$\begin{array}{ll} \text{minimize} & s_{\bar{\mathbf{z}}}(\mathbf{z}) \\ \text{subject to} & \mathbf{z} \in Z. \end{array} \tag{3.5.2}$$

Theorem 3.5.4. If the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is strictly increasing, then the solution of achievement problem (3.5.2) is weakly Pareto optimal. If the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is strongly increasing, then the solution of problem (3.5.2) is Pareto optimal. Finally, if the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is ε-strongly increasing, then the solution of problem (3.5.2) is ε-properly Pareto optimal.

Proof. Here we only prove the second statement because of the similarity of the proofs. We assume that $s_{\bar{\mathbf{z}}}$ is strongly increasing. Let $\mathbf{z}^* \in Z$ be a solution of the achievement problem. Let us suppose that it is not Pareto optimal. In this case, there exists an objective vector $\mathbf{z} \in Z$ such that $z_i \leq z_i^*$ for all $i = 1, \ldots, k$ and $z_j < z_j^*$ for some j. Because $s_{\bar{\mathbf{z}}}$ is strongly increasing, we know that $s_{\bar{\mathbf{z}}}(\mathbf{z}) < s_{\bar{\mathbf{z}}}(\mathbf{z}^*)$, which contradicts the assumption that $\mathbf{z}^*$ minimizes $s_{\bar{\mathbf{z}}}$. Thus $\mathbf{z}^*$ is Pareto optimal. □

The results of Theorem 3.5.4 can be augmented by the following theorem.

Theorem 3.5.5. If the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is increasing and the solution of achievement problem (3.5.2) is unique, then it is Pareto optimal.

Proof. The proof corresponds to the proof of Theorem 3.5.4.

Note that Theorems 3.5.4 and 3.5.5 are valid for any scalarizing function. Thus, the Pareto optimality and the weak Pareto optimality results proved for the weighting method, the ε-constraint method and the method of weighted metrics are explained by the monotonicity properties of the scalarizing functions in question (see, e.g., Vanderpooten (1990)).

We can now rewrite Theorem 3.5.4 so as to be able to characterize Pareto optimal solutions with the help of order-representing and order-approximating achievement functions. The proof follows from the proof of Theorem 3.5.4.

Corollary 3.5.6. If the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is order-representing, then, for any $\bar{\mathbf{z}} \in \mathbf{R}^k$, the solution of achievement problem (3.5.2) is weakly Pareto optimal. If the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is order-approximating with some ε and $\bar{\varepsilon}$ as in Definition 3.5.2, then, for any $\bar{\mathbf{z}} \in \mathbf{R}^k$, the solution of problem (3.5.2) is Pareto optimal. If $s_{\bar{\mathbf{z}}}$ in addition is $\bar{\varepsilon}$-strongly increasing, then the solution of problem (3.5.2) is $\bar{\varepsilon}$-properly Pareto optimal.

The preceding corollary states the sufficient conditions for a solution of an achievement function to be weakly, ε-properly, or Pareto optimal. The following theorem gives the corresponding necessary conditions.

Theorem 3.5.7. If the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is order-representing and $\mathbf{z}^* \in Z$ is weakly Pareto optimal or Pareto optimal, then it is a solution of achievement problem (3.5.2) with $\bar{\mathbf{z}} = \mathbf{z}^*$ and the value of the achievement function is zero. If the achievement function $s_{\bar{\mathbf{z}}}\colon Z \to \mathbf{R}$ is order-approximating and $\mathbf{z}^* \in Z$ is ε-properly Pareto optimal, then it is a solution of problem (3.5.2) with $\bar{\mathbf{z}} = \mathbf{z}^*$ and the value of the achievement function is zero.

Proof. Here, we only prove the statement for Pareto optimality. The proofs of the other statements are very similar. (The proof of the necessary condition for ε-proper Pareto optimality can be found in Wierzbicki (1986a).)

Let $\mathbf{z}^* \in Z$ be Pareto optimal. This means that there does not exist any other point $\mathbf{z} \in Z$ such that $z_i \leq z_i^*$ for all $i = 1, \ldots, k$ and $z_j < z_j^*$ for some j. Let us assume that $\mathbf{z}^*$ is not a solution of the achievement problem when $\bar{\mathbf{z}} = \mathbf{z}^*$. In this case there exists some vector $\mathbf{z}^\circ \in Z$ such that $s_{\bar{\mathbf{z}}}(\mathbf{z}^\circ) < s_{\bar{\mathbf{z}}}(\mathbf{z}^*) = s_{\bar{\mathbf{z}}}(\bar{\mathbf{z}}) = 0$ and $\mathbf{z}^\circ \neq \mathbf{z}^*$. Since $s_{\bar{\mathbf{z}}}$ was assumed to be order-representing, we have $\mathbf{z}^\circ \in \bar{\mathbf{z}} - \operatorname{int} \mathbf{R}_+^k = \mathbf{z}^* - \operatorname{int} \mathbf{R}_+^k$. This means that $z_i^\circ < z_i^*$ for all $i = 1, \ldots, k$, which contradicts the assumption that $\mathbf{z}^*$ is Pareto optimal. Thus, $\mathbf{z}^*$ is a solution of the achievement problem. □

Remark 3.5.8. Aided by the results in Theorem 3.5.7 a certain point can be confirmed not to be weakly, ε-properly or Pareto optimal (if the optimal value of the achievement function differs from zero).

We are now able to completely characterize the set of weakly Pareto optimal solutions with the help of order-representing achievement functions. The sets of Pareto optimal and ε-properly Pareto optimal solutions are characterized almost completely (if the closure of the sets of solutions of achievement problem (3.5.2) for an order-approximating achievement function is taken as $\varepsilon \to 0$). If the solutions of achievement problem (3.5.2) are assumed to be unique, the theorems above render the characterization of Pareto optimal solutions complete.

3.5.3. Comments

If the reference point is feasible, or to be more exact $\bar{\mathbf{z}} \in Z + \mathbf{R}_+^k$, then the minimization of the achievement function $s_{\bar{\mathbf{z}}}(\mathbf{z})$ subject to $\mathbf{z} \in Z$ must produce a solution that maximizes the distance to the Pareto optimal set. If the reference point is infeasible, that is, $\bar{\mathbf{z}} \notin Z + \mathbf{R}_+^k$, then the minimization of the achievement function $s_{\bar{\mathbf{z}}}(\mathbf{z})$ subject to $\mathbf{z} \in Z$ must produce a solution that minimizes the distance to the Pareto optimal set.

The advantage of the achievement functions is that any arbitrary weakly Pareto optimal or Pareto optimal (or at least ε-properly Pareto optimal) solution can be obtained by moving the reference point only. It is shown in Wierzbicki (1986a) that the solution of the achievement function depends Lipschitz-continuously on the reference point.

There are many achievement functions satisfying the above-presented conditions. An example of order-representing functions was given in problem (3.5.1). The corresponding achievement function is

$$s_{\bar{\mathbf{z}}}(\mathbf{z}) = \max_{i=1,\dots,k} [\, w_i(z_i - \bar{z}_i) \,],$$

where $\mathbf{w}$ is some fixed positive weighting vector. Let us briefly convince ourselves that the above-mentioned function really is order-representing. The continuity of the function is obvious. If we have $\mathbf{z}^1$ and $\mathbf{z}^2 \in Z$ such that $z_i^1 < z_i^2$ for all $i = 1, \dots, k$, then $s_{\bar{\mathbf{z}}}(\mathbf{z}^1) = \max_i [w_i(z_i^1 - \bar{z}_i)] < \max_i [w_i(z_i^2 - \bar{z}_i)] = s_{\bar{\mathbf{z}}}(\mathbf{z}^2)$ and thus the function is strictly increasing. If the inequality $s_{\bar{\mathbf{z}}}(\mathbf{z}) = \max_i [w_i(z_i - \bar{z}_i)] < 0$ holds, then we must have $z_i < \bar{z}_i$ for all $i = 1, \dots, k$, that is, $\mathbf{z} \in \bar{\mathbf{z}} - \operatorname{int} \mathbf{R}_+^k$.

An example of order-approximating achievement functions is

$$s_{\bar{\mathbf{z}}}(\mathbf{z}) = \max_{i=1,\dots,k} [\, w_i(z_i - \bar{z}_i) \,] + \rho \sum_{i=1}^{k} w_i(z_i - \bar{z}_i), \tag{3.5.3}$$

where $\mathbf{w}$ is some fixed positive weighting vector and $\rho > 0$ is sufficiently small when compared with ε and large when compared with $\bar{\varepsilon}$. The weighting coefficients can also be dropped from the latter part. This function is also $\bar{\varepsilon}$-strongly increasing. Function (3.5.3) is related to augmented weighted Tchebycheff problem (3.4.5) and, thus, it can be called an augmented weighted achievement function.

An example of a so-called *penalty scalarizing function* is

$$s_{\bar{\mathbf{z}}}(\mathbf{z}) = -\|\mathbf{z} - \bar{\mathbf{z}}\|^2 + \varrho \|(\mathbf{z} - \bar{\mathbf{z}})_+\|^2,$$

where $\varrho > 1$ is a scalar penalty coefficient and $(\mathbf{z} - \bar{\mathbf{z}})_+$ is a vector with components $\max [0, z_i - \bar{z}_i]$. This function is strictly increasing, strongly increasing for all the metrics in $\mathbf{R}^k$ except for the Tchebycheff metric and order-approximating with $\varepsilon \geq 1/\varrho$ (see Wierzbicki (1980a, 1982)). More examples of order-representing and order-approximating functions are presented, for example, in Wierzbicki (1980b, 1986a, b).

In cases when there exists a weighting vector such that the solution of weighting problem (3.1.1) is equal to the solution of the achievement problem, the weighting vector can be obtained from partial derivative information of the achievement function. See Wierzbicki (1982) for details.

Let us finally mention a subset of reference points, termed dominating points, considered in Skulimowski (1989). A point is called a *dominating point*

if it is not dominated by any feasible point and it dominates at least one of the feasible points.

3.5.4. Concluding Remarks

Achievement scalarizing functions are a set of general scalarizing functions satisfying certain requirements. In general, achievement functions are conceptually very appealing for generating weakly, ε-properly or Pareto optimal solutions. They overcome most of the difficulties arising with other methods in this class.

The results concerning achievement functions will be utilized, for example, in Section 5.6 when deriving an interactive method. For interactive methods, the idea of moving the reference point instead of the weighting coefficient seems more natural and easier for the decision maker. A fact favouring achievement scalarizing functions against weighted metrics is that the global ideal objective vector does not have to be known. Thus, the method is more reliable.

3.6. Other A Posteriori Methods

Finally, we briefly mention some other methods of the a posteriori type. For more detailed information, see the references cited.

The so-called hyperplane method is introduced in Yano and Sakawa (1989) for generating Pareto optimal or properly Pareto optimal solutions. It is shown that the weighting method, the ε-constraint method and the method of weighted metrics can be viewed as special cases of the hyperplane method. A theory concerning trade-off rates in the hyperplane method is provided in Sakawa and Yano (1990). A generalized hyperplane method for generating all the efficient solutions (with respect to some ordering cone) is presented in Sakawa and Yano (1992).

Another method for a general characterization of the Pareto optimal set is suggested in Soland (1979). For example, the weighting method, the method of weighted metrics and goal programming (see Section 4.3) can be seen as special cases of the general scalar problem of Soland. Further, the weighting method and the ε-constraint method are utilized in a so-called envelope approach for determining Pareto optimal solutions in Li and Haimes (1987). An application to dynamic multiobjective programming is also treated.

The noninferior (meaning here Pareto optimality) set estimation (NISE) method for MOLP problems can also be considered to belong to this class of a posteriori methods. It is a technique for generating the Pareto optimal set of two objective functions (see Cohon (1978)). It can be generalized for convex problems with two objective functions (see, for example, Chankong and Haimes (1983b, pp. 268–274)). In Balachandran and Gero (1985), the method is extended to problems with three objective functions. The weighting method is the basis of the NISE method.

Multiobjective optimization problems with polynomial objective and constraint functions are treated in Kostreva et al. (1992). The method for determining Pareto optimal solutions is based on problem (2.10.2) of Part I and a so-called homotopy continuation. Note that problems with polynomial functions are highly nonlinear, nonconvex and nonconcave.

A scalarization method for multiobjective optimization problems, where optimality is defined through ordering cones, is suggested in Pascoletti and Serafini (1984). By varying the parameters of the scalar problem it is possible to find all the efficient solutions. A further investigation is conducted in Sterna-Karwat (1987).

It is suggested in Benson and Sayin (1997) that instead of trying to generate the whole Pareto optimal set one should aim at finding a truly global representation of it. This would decrease both the burden of the decision maker and computational costs. Benson and Sayin introduce a global shooting procedure to meet this need. In Armann (1989), a method is presented for generating a dispersed subset of the Pareto optimal set, which is then presented to the decision maker.

One more method for generating an evenly distributed set of Pareto optimal solutions to a differentiable nonlinear multiobjective optimization problem is suggested in Das and Dennis (1998). The approach is called the normal boundary intersection (NBI) method. The idea in broad outline is to intersect the feasible objective region with a normal to the convex combinations of the columns of the payoff matrix. Evenly distributed parameters, that is, the coefficients in the convex combinations, produce evenly distributed solutions. The weakness of the approach is the fact that it may produce non-Pareto optimal solutions to nonconvex problems.

The difficulty of illustrating the set of Pareto optimal solutions to the decision maker is treated in Bushenkov et al. (1994, 1995) and Lotov et al. (1992, 1997). An extension of the Pareto optimal set, a so-called Pareto optimal hull, is approximated by polyhedral sets (see Lotov (1995, 1996)) using convolution-based algorithms (see Bushenkov et al. (1995) and Lotov (1996)). Different decision maps are generated in this way (see Chapter 3 of Part III). Specific approaches exist for linear, convex and nonconvex cases but they are all based on the same so-called generalized reachable sets method. An implementation of the generalized reachable sets method is available (see Section 2.2 in Part III).

4. A PRIORI METHODS

In the case of a priori methods, the decision maker must specify her or his preferences, hopes and opinions before the solution process. The difficulty is that the decision maker does not necessarily know beforehand what it is possible to attain in the problem and how realistic her or his expectations are. The working order in these methods is: 1) decision maker, 2) analyst.

Below, we handle three a priori methods. First, we give a short presentation of the value function method. Then we introduce lexicographic ordering and goal programming.

4.1. Value Function Method

The value function optimization approach was already mentioned earlier. Here we present it again briefly.

4.1.1. Introduction

In the value function method, the decision maker must be able to give an accurate and explicit mathematical form of the value function $U: \mathbf{R}^k \to \mathbf{R}$ that represents her or his preferences globally. This function provides a complete ordering in the objective space. Then the *value function problem*

$$\begin{aligned} &\text{maximize} && U(\mathbf{f}(\mathbf{x})) \\ &\text{subject to} && \mathbf{x} \in S \end{aligned} \tag{4.1.1}$$

is ready to be solved by some method for single objective optimization as illustrated in Figure 4.1.1. The bold line represents the Pareto optimal set. Remember Theorem 2.6.2 of Part I, which says that the solution of problem (4.1.1) is Pareto optimal if the value function is strongly decreasing.

The value function method seems to be a very simple method, but the difficulty lies in specifying the mathematical expression of the value function. The inability to encode the decision maker's underlying value function reliably is demonstrated by experiments in de Neufville and McCord (1984). It is shown that encoding methods that should theoretically produce identical value functions fail: the functions may differ from each other by more than 50 %. It is also

K. Miettinen, *Nonlinear Multiobjective Optimization*

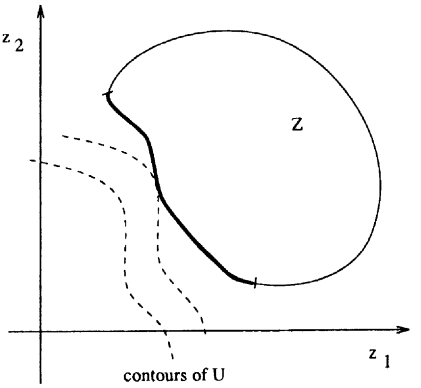

Figure 4.1.1. Contours of the value function.

pointed out that there is no actual analysis of the accuracy of the value function assessment. The consistency checks, that is, whether decision makers provide consistent answers to similar questions, are not adequate: a biased instrument can provide consistent data.

On the other hand, even if it were possible for the decision maker to express her or his preferences globally, the resulting preference structure might be too simple, since value functions cannot represent intransitivity or incomparability (see Rosinger (1985)). More features and weaknesses were presented in connection with the definition of the value function (Definition 2.6.1) in Section 2.6 of Part I.

4.1.2. Comments

The value function method could be called an 'optimal' way of solving multiobjective optimization problems if the decision maker could reliably express the value function. The use of the value function method is restricted in practice to multiattribute decision analysis problems with a discrete set of feasible alternatives. The theory of value and utility functions for multiattribute problems is examined broadly in Keeney and Raiffa (1976). But, it is believed, for example, in Rosenthal (1985), that these experiences can also be utilized in continuous cases.

Important results concerning value functions and the conditions for their existence are collected in Dyer and Sarin (1981). Two general classes of value functions, additive and multiplicative forms, are presented extensively in Keeney and Raiffa (1976) and briefly in Rosenthal (1985). The existence of value functions and the nature of additive decreasing value functions are handled in Stam et al. (1985). These topics and the construction of value functions are presented more widely in Yu (1985, pp. 95–161). General properties and some desirable features of certain types of value functions (e.g., additive, max-min,

min-sum and exponential forms) are stated in Bell (1986), Harrison and Rosenthal (1988), Soland (1979) and Sounderpandian (1991). More examples of value functions are given in Tell and Wallenius (1979). Utility compatible measures of risk are deduced in Bell (1995). Relations between value functions, ordering cones and (proper) efficiency are studied in Henig (1990).

In some interactive methods, it is assumed that the underlying value function is of some particular (e.g., additive or exponential) form, after which, its parameters are fitted according to the decision maker's preferences. Such methods are presented, for example, in Rothermel and Schilling (1986) and Sakawa and Seo (1980, 1982a, b) (see Section 5.3).

Three kinds of conditions for value functions under which it is not possible to exclude any Pareto optimal or properly Pareto optimal solution from consideration a priori are identified in Soland (1979).

The convergence properties of additive value functions (assuming preferential independence of the objective functions) are investigated by simulation experiments in Stewart (1997). One observation is that piecewise linear value functions perform dramatically better than linear ones.

Relationships between the method of weighted metrics and the value function method are reported in Ballestero and Romero (1991). It might be imagined that the two methods have nothing in common, since a value function represents the opinions of the decision maker and the method of weighted metrics does not take the decision maker into consideration. However, conditions can be set on the value function to guarantee that its optimum belongs to the solution set obtainable by the method of weighted metrics. More relationships between these two methods, when the value functions are of a certain type, are presented in Ballestero (1997a). It is demonstrated in Morón et al. (1996) that there are large families of such well-behaved value functions for bi-criteria problems where the connection is valid.

4.1.3. Concluding Remarks

The value function method is an excellent method if the decision maker happens to know an explicit mathematical formulation for the value function and if that function represents wholly the preferences of the decision maker. These two crucial preconditions are the difficulties of the approach.

There are certain conditions that the decision maker's preferences must satisfy so that a value function can be defined on them. The decision maker must, for instance, be able to specify consistent (implying transitive) preferences. Thus, there may not necessarily exist a value function that will impose a total order in the set of feasible objective vectors. The assumption of a total order is often contrary to our intuitive aims and hence is quite likely to lead to less than ideal selections, as Polak and Payne (1976) remind us. This fact must be kept in mind below, when several methods which assume the existence of a value function (at least implicitly) are introduced.

One important thing to take into account in practice is that the aspirations of the decision maker may change during the solution process. Possible explanations of such behaviour are pondered in Steuer and Gardiner (1990). Is it possible that the decision maker's value function will change considerably over a short period of time and thus be unstable? Another alternative is that it is difficult for the decision maker to know the real value function without getting to know the problem better, that is, without interaction with the solution process. More open questions concerning value functions are listed in Nijkamp et al. (1988).

The weighting method may be regarded as a special case of a value function where the utilities are linear and additive. If the underlying value function is assumed to be linear, this means that the marginal rates of substitution of the decision maker are constant for every solution. See comments on this feature in Section 4.3.

4.2. Lexicographic Ordering

Lexicographic ordering was mentioned earlier as a tool for producing Pareto optimal solutions from weakly Pareto optimal ones. It can also be used as an a priori solution method.

4.2.1. Introduction

In lexicographic ordering the decision maker must arrange the objective functions according to their absolute importance. This ordering means that a more important objective is infinitely more important than a less important objective. After ordering, the most important objective function is minimized subject to the original constraints. If this problem has a unique solution, it is the solution of the whole multiobjective optimization problem. Otherwise, the second most important objective function is minimized. Now, in addition to the original constraints, a new constraint is added. This new constraint is there to guarantee that the most important objective function preserves its optimal value. If this problem has a unique solution, it is the solution of the original problem. Otherwise, the process goes on as above. Lexicographic orders and utilities are widely examined in Fishburn (1974).

An example of lexicographic ordering is presented in Figure 4.2.1. There are two objective functions of which the first is the most important. After minimizing the first objective, there are two points left and after minimizing the second objective, the point $\mathbf{z}^1$ is obtained as the final solution. The bold line represents the Pareto optimal set in the figure. This example is somewhat too positive since all the objective functions have their effect on the solution process.

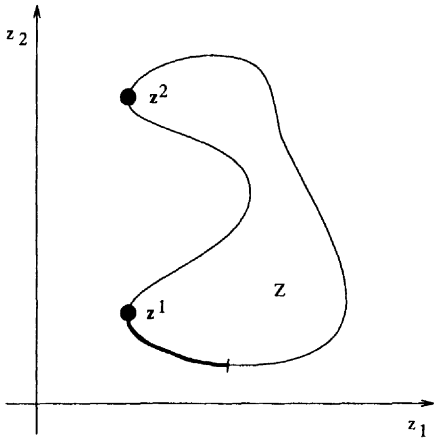

Figure 4.2.1. Lexicographic ordering.

Let the objective functions be arranged according to the lexicographic order from the most important f_1 to the least important f_k. We write the *lexicographic problem* as

$$\begin{aligned} &\text{lex minimize} && f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}) \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{4.2.1}$$

We can now present the following result concerning the Pareto optimality of the solutions.

Theorem 4.2.1. The solution of lexicographic problem (4.2.1) is Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a solution of the lexicographic problem. Let us assume that it is not Pareto optimal. In this case, there exists a point $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and for at least one j the inequality is strict, that is, $f_j(\mathbf{x}) < f_j(\mathbf{x}^*)$.

Let be $i = 1$. From the definition of lexicographic ordering we know that f_1 attains its minimum at $\mathbf{x}^*$. Since also $f_1(\mathbf{x}) \leq f_1(\mathbf{x}^*)$, it is only possible that $f_1(\mathbf{x}) = f_1(\mathbf{x}^*)$.

There are two possibilities in determining the lexicographic optimum. Either a unique solution is found during the optimization process, or optimizations are performed for every $i = 1, \ldots, k$. In the latter case, where $i = 2$, we also have $f_2(\mathbf{x}) = f_2(\mathbf{x}^*)$ and with similar reasoning we have that $f_i(\mathbf{x}) = f_i(\mathbf{x}^*)$ for every $i = 1, \ldots, k$. This contradicts the assumption of at least one strict inequality. Thus, $\mathbf{x}^*$ is Pareto optimal.

On the other hand, if lexicographic ordering stops before every objective function has been examined, this means that a unique solution $\mathbf{x}^*$ has been

obtained for f_i. The assumption $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ implies that $f_i(\mathbf{x}) = f_i(\mathbf{x}^*)$, which is a contradiction. Thus, $\mathbf{x}^*$ is Pareto optimal. □

4.2.2. Comments

Numerical application examples of the method are given in Hwang and Masud (1979, pp. 49–55). In Ben-Tal (1980), Pareto and lexicographic optima are characterized in convex problems. Duality theory for convex problems with the help of lexicographic ordering is developed in Martínez-Legaz (1988).

A modification of lexicographic ordering, called hierarchical optimization, is applied to a vehicle design problem of mechanical engineering in Bestle and Eberhard (1997). In hierarchical optimization the upper bounds obtained when minimizing more important objective functions are relaxed by so-called worsening factors. These factors are specified by the decision maker.

Lexicographic ordering corresponds to the weighting method when the weighting coefficients are of very different magnitude. The question whether there exist weighting vectors such that the optimal solution of the weighting method is identical to the solution obtained by lexicographic ordering is considered in Sherali (1982) and Sherali and Soyster (1983). The answer is positive for linear problems and several discrete problems. In practice, this means that the problem of lexicographic ordering can be solved as a weighting problem with standard optimizers.

The notion absolute importance of objective functions is discussed in Roy and Mousseau (1996). Roy and Mousseau also consider under what kind of conditions one can say that one objective function is more important than another.

4.2.3. Concluding Remarks

The justification for using lexicographic ordering is its simplicity and the fact that people usually make decisions successively. However, this method has several drawbacks. The decision maker may have difficulties in putting the objective functions into an absolute order of importance. On the other hand, the method is usually robust. It is very likely that the less important objective functions are not taken into consideration at all. If the most important objective function has a unique solution, the other objectives do not have any influence on the solution. And even if the most important objective had alternative optima and it was possible to use the second most important objective, it is very unlikely that this problem would have alternative optima, and the third or other less important objectives could be used.

Note that lexicographic ordering does not allow a small increment of an important objective function to be traded off with a great decrement of a less important objective function. Yet, this kind of trading might often be appealing to the decision maker.

Lexicographic ordering may be used as a part of the following solution method, called goal programming.

4.3. Goal Programming

The ideas of goal programming were originally introduced in Charnes et al. (1955), but the term goal programming was fixed in Charnes and Cooper (1961). It is one of the first methods expressly created for multiobjective optimization. Among more recent papers, an easy-to-understand presentation of goal programming is given in Ignizio (1983a, 1985). Goal programming was originally developed for MOLP problems, and this background is very evident in the formulation.

4.3.1. Introduction

The basic idea in goal programming is that the decision maker specifies (optimistic) aspiration levels for the objective functions and any deviations from these aspiration levels are minimized. An objective function jointly with an aspiration level forms a *goal*. We can say that, for example, minimizing the price of a product is an objective function, but if we want the price to be less than 500 dollars, it is a goal (and if the price must be less than 500 dollars, it is a constraint). We denote the aspiration level of the objective function f_i by $\bar{z}_i$ for $i = 1, \ldots, k$.

For minimization problems, goals are of the form $f_i(\mathbf{x}) \leq \bar{z}_i$ (and of the form $f_i(\mathbf{x}) \geq \bar{z}_i$ for maximization problems). Goals may also be represented as equalities or ranges (for the latter, see Charnes and Cooper (1977)). The aspiration levels are assumed to be selected so that they are not achievable simultaneously.

It is worth noticing that the goals are of the same form as the constraints of the problem. This is why the constraints may be regarded as a subset of the goals. This way of formulating the problem is called *generalized goal programming*. In this case, the goals can be thought of as being divided into *flexible* and *inflexible goals*, where the constraints are the inflexible (or rigid) ones. More detailed presentations and practical applications of generalized goal programming are given, for example, in Ignizio (1983a) and Korhonen (1991a). See also Section 5.10.

After the aspiration levels have been specified, the following task is to minimize the under- and overachievements of the objective function values with respect to the aspiration levels. It is sufficient to study the deviational variables $\delta_i = \bar{z}_i - f_i(\mathbf{x})$. The deviational variable δ_i may have positive or negative values, depending on the problem. We can present it as the difference of two positive variables, that is, $\delta_i = \delta_i^- - \delta_i^+$. We can now investigate how well each of the aspiration levels is attained by studying the deviational variables. We

can write $f_i(\mathbf{x}) + \delta_i^- - \delta_i^+ = \bar{z}_i$ for all $i = 1, \ldots, k$, where δ_i^- is a *negative deviation* or *underachievement* and δ_i^+ is a *positive deviation* or *overachievement* in relation to the aspiration level. It is valid that $\delta_i^- \cdot \delta_i^+ = 0$ for all $i = 1, \ldots, k$.

We now have the multiobjective optimization problem in a form where we can minimize the deviational variables. For minimization problems it is sufficient to minimize the k variables δ_i^+. If the ith goal is in the form of an equality, we minimize $\delta_i^- + \delta_i^+$.

4.3.2. Different Approaches

Thus far, we have only formulated the multiobjective optimization problem in an equivalent form, where we have deviational variables as the objective functions. There are several ways to proceed from this point. Here we present a weighted (also called Archimedian) and a lexicographic (also called preemptive) approach. More methods are handled in Ignizio (1983a) and some formulations are explored in de Kluyver (1979).

In the *weighted* approach, see Charnes and Cooper (1977), the weighted sum of the deviational variables is minimized. This means that in addition to the aspiration levels, the decision maker must specify information about the importance of attaining the aspiration levels in the form of weighting coefficients. The weighting coefficients are assumed to be positive and sum up to one. The bigger the weighting coefficient is, the more important is the attainment of that aspiration level. (Sometimes negative weighting coefficients are used to represent a premium instead of a penalty.)

To put the introduction presented above into mathematical form and to reason about the usage of the deviation variables, we can say that the problem

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{k} w_i |\bar{f}_i(\mathbf{x}) - z_i| \\ &\text{subject to} && \mathbf{x} \in S \end{aligned} \tag{4.3.1}$$

is converted into a new form by adding the overachievement variables

$$\delta_i^+ = \max[0,\ f_i(\mathbf{x}) - \bar{z}_i] \quad \text{or} \quad \delta_i^+ = \frac{1}{2}\Big[|\bar{z}_i - f_i(\mathbf{x})| + f_i(\mathbf{x}) - \bar{z}_i\Big]$$

and underachievement variables

$$\delta_i^- = \max[0,\ \bar{z}_i - f_i(\mathbf{x})] \quad \text{or} \quad \delta_i^- = \frac{1}{2}\Big[|\bar{z}_i - f_i(\mathbf{x})| + \bar{z}_i - f_i(\mathbf{x})\Big].$$

This means that the absolute value signs can be dropped from problem (4.3.1) by introducing the underachievement and the overachievement variables. The resulting *weighted goal programming problem* is

$$
\begin{aligned}
&\text{minimize} && \sum_{i=1}^{k}(w_i^- \delta_i^- + w_i^+ \delta_i^+) \\
&\text{subject to} && f_i(\mathbf{x}) + \delta_i^- - \delta_i^+ = \bar{z}_i \quad \text{for all} \quad i = 1, \ldots, k, \\
& && \delta_i^-, \delta_i^+ \geq 0 \quad \text{for all} \quad i = 1, \ldots, k, \\
& && \mathbf{x} \in S,
\end{aligned} \tag{4.3.2}
$$

where we give separate weighting coefficients for underachievements and overachievements, and $\mathbf{x} \in \mathbf{R}^n$, δ_i^- and δ_i^+, $i = 1, \ldots, k$, are the variables. If all the goals are in the form $f_i(\mathbf{x}) \leq \bar{z}_i$, we can leave the underachievement variables and write the problem in the form

$$
\begin{aligned}
&\text{minimize} && \sum_{i=1}^{k} w_i^+ \delta_i^+ \\
&\text{subject to} && f_i(\mathbf{x}) - \delta_i^+ \leq \bar{z}_i \quad \text{for all} \quad i = 1, \ldots, k, \\
& && \delta_i^+ \geq 0 \quad \text{for all} \quad i = 1, \ldots, k, \\
& && \mathbf{x} \in S,
\end{aligned} \tag{4.3.3}
$$

where $\mathbf{x} \in \mathbf{R}^n$ and δ_i^+, $i = 1, \ldots, k$, are the variables.

Figure 4.3.1 portrays how problem (4.3.3) is solved. The black spot is the reference point of the aspiration levels. Every weighting vector produces different contours by which the feasible objective region is to be intersected. Thus, different solutions can be obtained by altering the weights. Contours with two weighting vectors have been depicted in the figure. The bold line illustrates the Pareto optimal set.

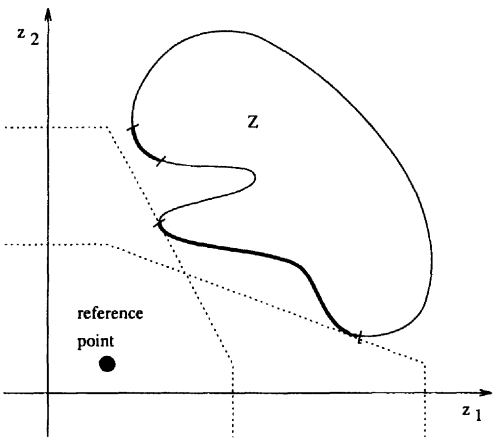

Figure 4.3.1. Contours with different weighting vectors.

Even though the constraints $\delta_i^- \cdot \delta_i^+ = 0$ for all $i = 1, \ldots, k$ are not usually included in the problem formulations, some attention must be paid to guar-

antee that they are valid (see details in Rosenthal (1983)). An example of the required conditions is given in Sawaragi et al. (1985, p. 253). The weighted goal programming problem may be solved by standard single objective optimization methods. If the original problem is linear, then the corresponding weighted goal programming problem is also linear. The close connection between goal programming and MOLP problems explains why the above-mentioned constraint is usually absent from the problem formulation (it would make the problem nonlinear).

Note that weighted goal programming is closely related to the method of weighted metrics or compromise programming. This can be seen particularly well in formulation (4.3.1). Instead of the ideal objective vector, the reference point of the decision maker is used in goal programming. The distances can be measured by metrics other than the L_1-metric. The L_1-metric is widely used in connection with goal programming because of the origin of the method in linear programming. (This metric maintains the linearity of the problem.) If some other L_p-metric is used there is another problem in determining an appropriate value for p. Note, however, that if we have appropriate solvers available, we can solve problem (4.3.1) directly without any deviational variables and using any metric.

In the *lexicographic* approach, the decision maker must specify a lexicographic order for the goals in addition to the aspiration levels. The goal at the highest priority level is supposed to be infinitely more important than the goal at the second priority level, etc. This means that no matter how large a multiplier is selected, a lower priority goal multiplied by it can never be made as important as a higher priority goal. After the lexicographic ordering, the problem with the deviational variables as objective functions and the constraints as in (4.3.2) is solved as explained in Section 4.2. In order to be able to use the lexicographic approach, the decision maker's preference order for the objectives must be definite and rigid.

A combination of the weighted and the lexicographic approaches, to be called a *combined* approach, is quite popular. In this case, several objective functions may belong to the same class of importance in the lexicographic order. In each priority class, a weighted sum of the deviational variables is minimized. The same weaknesses presented in connection with lexicographic ordering are valid for this and the lexicographic approach.

It is not necessary to include the original constraints ($\mathbf{x} \in S$) in the lexicographic optimization problem in the normal way. They can be considered to belong to the first priority level. In this way, they are taken into account before any objective function is optimized and the feasibility of the solutions is guaranteed by the nature of the lexicographic ordering.

Next, we prove a result concerning the Pareto optimality of the solutions of goal programming.

Theorem 4.3.1. The solution of a weighted or a lexicographic goal programming problem is Pareto optimal if either the aspiration levels form a Pareto optimal reference point or all the deviational variables δ_i^+ for functions to be minimized and δ_i^- for functions to be maximized have positive values at the optimum.

Proof. For the lexicographic approach, the proof corresponds to that of Theorem 4.2.1. Here, we only present a proof for the weighted approach. For simplicity of notation, we assume that the problem is of the form (4.3.3). A more general case is straightforward.

Let $\mathbf{x}^* \in S$ be a solution of the weighted goal programming problem, where the deviational variables (denoted here for clarity by δ_i^*) are positive. Let us assume that $\mathbf{x}^*$ is not Pareto optimal. In this case, there exists a vector $\mathbf{x}^\circ \in S$ such that $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*)$ for at least one index j.

We denote $f_j(\mathbf{x}^*) - f_j(\mathbf{x}^\circ) = \beta > 0$. Then we set $\delta_i^\circ = \delta_i^* > 0$ for $i \neq j$ and $\delta_j^\circ = \max[0, \delta_j^* - \beta] \geq 0$, where δ_i° is the deviational variable corresponding to $\mathbf{x}^\circ$ for $i = 1, \ldots, k$.

We have now $f_i(\mathbf{x}^\circ) - \delta_i^\circ \leq f_i(\mathbf{x}^*) - \delta_i^* \leq \bar{z}_i$ for all $i \neq j$. If $\delta_j^* - \beta > 0$, then $f_j(\mathbf{x}^\circ) - \delta_j^\circ = f_j(\mathbf{x}^\circ) - \delta_j^* + f_j(\mathbf{x}^*) - f_j(\mathbf{x}^\circ) \leq \bar{z}_j$, and if $\delta_j^* - \beta \leq 0$, then $f_j(\mathbf{x}^\circ) - \delta_j^\circ = f_j(\mathbf{x}^\circ) + f_j(\mathbf{x}^*) - f_j(\mathbf{x}^*) = f_j(\mathbf{x}^*) - \beta \leq f_j(\mathbf{x}^*) - \delta_j^* \leq \bar{z}_j$.

This means that $\mathbf{x}^\circ$ satisfies the constraints of problem (4.3.3). We have $\delta_j^\circ < \delta_j^*$ (this is also valid if $\delta_j^\circ = 0$ since $\delta_i^* > 0$ for all i), and $\delta_i^\circ \leq \delta_i^*$ for all $i \neq j$. As the weighting coefficients are positive, we have $\sum w_i^+ \delta_i^\circ < \sum w_i^+ \delta_i^*$, which contradicts the fact that $\mathbf{x}^*$ is a solution of weighted goal programming problem (4.3.3).

For aspiration levels forming a Pareto optimal point the proof is self-evident. □

Let us briefly mention one more form of goal programming, *min-max goal programming* (suggested in Flavell (1976)). It is not as widely used as the weighted and the lexicographic approaches. For minimization problems the *min-max goal programming problem* to be solved is

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} \delta_i^+ \\ &\text{subject to} && f_i(\mathbf{x}) - \delta_i^+ \leq \bar{z}_i \quad \text{for all} \quad i = 1, \ldots, k, \\ & && \mathbf{x} \in S, \end{aligned} \tag{4.3.4}$$

where $\mathbf{x} \in \mathbf{R}^n$ and δ_i^+, $i = 1, \ldots, k$, are the variables.

4.3.3. Comments

If the optimal objective function value of the goal programming problem equals zero, some caution is in order, since the solution obtained may not be Pareto optimal. The reason is that if the aspiration levels are all feasible, then the value zero for all the deviational variables gives the minimum value (zero) for the goal programming objective function. Thus the solution is equal to the reference point, and normally there exist many feasible points that are not Pareto optimal. If the solutions are intended to be Pareto optimal despite the selection of the aspiration levels, then we must maximize the distance if the aspiration levels are feasible and minimize the distance if the aspiration levels are infeasible. This is the case with achievement scalarizing functions, as explained in Section 3.5.

It is shown in Caballero et al. (1996) that the solution of min-max goal programming problem (4.3.4) is Pareto optimal if it is unique. Other tests for Pareto optimality in goal programming are provided in Romero (1991).

It is pointed out in Romero (1997) that lexicographic orderings imply discontinuous preferences. This means that lexicographic goal programming is incompatible with ordering the decision maker's preferences by a decreasing value function. Thus, Romero recommends caution with lexicographic goal programming because it is applicable only for problems with discontinuous preferences. In Romero (1997), the importance of knowing the different preferential logics underlying each goal programming approach is also emphasized. Different logics are introduced and the min-max approach is slightly favoured, but mixtures of the approaches are recommended.

The lexicographic goal programming approach can be modified so that it can even take into account goals with lower priority, as suggested in Caballero et al. (1996, 1997). It is proved in Caballero et al. (1996) that the solution of the modified lexicographic goal programming problem is Pareto optimal if the solution of the lowest priority level (i.e., the last optimization problem) is unique.

As pointed out in Dyer and Sarin (1981), although it is not readily apparent, goal programming implicitly assumes that there is a measurable, additive and rigid piecewise linear underlying value function. Rosenthal stresses, in Rosenthal (1983), that weighted goal programming problem (4.3.2) is equivalent to the value function maximization problem where

$$\frac{\partial U(\mathbf{f}(\mathbf{x}))}{\partial f_i} = \begin{cases} w_i^- & \text{if } f_i(\mathbf{x}) < \bar{z}_i, \\ -w_i^+ & \text{if } f_i(\mathbf{x}) > \bar{z}_i, \end{cases}$$

which means that the marginal utility is constant on either side of the aspiration level. This is contrary to the economic idea that a decision maker considers the next unit of decrease of f_i more important when f_i is plentiful than when f_i is scarce.

This idea is even more evident when we look at the marginal rates of substitution in goal programming problems. In Remark 2.8.7 of Part I it was mentioned that the marginal rates of substitution may be defined as $m_{ij}(\mathbf{x}) = \frac{\partial U(\mathbf{f}(\mathbf{x}))}{\partial f_j} / \frac{\partial U(\mathbf{f}(\mathbf{x}))}{\partial f_i}$. Thus, goal programming does not take into consideration the possibility that it is easier for the decision maker to let something increase a little if (s)he has got little of it than if (s)he has got much of it. The reason for this is that goal programming implicitly assumes that the marginal rates of substitution are piecewise constant. This critique also applies to the lexicographic approach (see details in Rosenthal (1983, 1985)). More critical observations about goal programming are presented in Romero (1991) and Rosenthal (1983).

4.3.4. Applications and Extensions

A comprehensive presentation on goal programming and its extensions is given in Ignizio (1976), and a summary of different variations of goal programming is provided in Charnes and Cooper (1977). In addition, a wide survey of the literature around goal programming up to the year 1983 is presented in Soyibo (1985). Several modifications and improvements as well as applications are reviewed. A survey of goal programming is also given in Kornbluth (1973) and the weighted and the lexicographic approaches are applied to problems with fractional objective functions. Further, a broad collection of journal papers and books on goal programming is assembled in Schniederjans (1995a). References in nine broad areas of application are also included.

In the literature, goal programming is the most widely used solution method in terms of practical applications. Weighted goal programming with equal weighting coefficients is employed in the planning of public works in Yoshikawa et al. (1982). Weighted goal programming with sensitivity analysis is also used for portfolio selection in Tamiz et al. (1996).

Lexicographic goal programming is applied in Benito-Alonso and Devaux (1981) to a problem concerning the location and size of day nurseries, in Sinha et al. (1988) to storage problems in agriculture and in Mitra and Patankar (1990) to aid manufacturers in selecting the price and the warranty time of their products. Lexicographic goal programming is also applied in Kumar et al. (1991) to nonlinear multi-stage decision problems in manufacturing systems, in Ng (1992) to aircraft loading and in Brauer and Naadimuthu (1992) to solve a mixed integer MOLP problem involving inventory and distribution planning. In Hemaida and Kwak (1994) a linear trans-shipment problem and in Current and Storbeck (1994) a location model are solved by lexicographic goal programming, and in Giannikos et al. (1995) it is applied in an integer allocation problem. In Berbel and Zamora (1996) lexicographic goal programming is applied in wildlife management and in Kim et al. (1997) in solving a linear problem of military budget planning. An implementing decision support system is also described.

A numerical application example of combined goal programming is given in Hwang and Masud (1979, pp. 79–95). Combined goal programming is applied

in Levary (1986) to problems of optimal control, in Giokas and Vassiloglou (1991) to the (linear) management of the assets and liabilities of a Greek bank and in Ghosh et al. (1992) to the resource planning of university management. In Sankaran (1990), the combined approach is also used to solve an integer MOLP problem in cell formation, and in Schniederjans and Hoffman (1992), combined zero-one goal programming is applied to a problem concerning international business expansion analysis. The ideas of combined goal programming are adapted in Miyaji et al. (1988) in solving a transportation problem-type problem of dividing students into groups. In addition, the combined goal programming approach is applied in fund and portfolio management in Powell and Premachandra (1998).

The applications mentioned here are only a few of the existing ones. The popularity of goal programming is well affirmed by the fact that in a bibliography collected in White (1990) on multiobjective optimization applications (covering the years from 1955 to 1986) more than a half involved goal programming.

Four different goal interpretations in multiobjective optimization are presented in Dinkelbach (1980). Goal programming is adapted to multiobjective generalized networks for integer problems in Ignizio (1983b). In Inuiguchi and Kume (1991), goal programming is extended to linear problems where the coefficients and the aspiration levels are given as intervals. The aspiration level intervals do not there represent regions within which the decision maker is satisfied, but regions where the aspiration levels may vary. A generalization of goal programming through the theory of variational inequalities is presented in Thore et al. (1992).

An extension of goal programming to MOLP problems is given in Martel and Aouni (1990). Instead of the deviational variables, some functions describing the wishes of the decision maker about attaining the goals set are used in the weighted approach. An illustrative example is also provided. Technical improvements to the preference modelling method of Martel and Aouni are presented in Tamiz and Jones (1995). This approach is extended in Martel and Aouni (1998) by allowing goals to be intervals instead of exact numbers. This means that indifference thresholds (see Subsection 5.9.1) are used in modelling the imprecision of the goals. (Even though we have mentioned some interesting MOLP extensions and solution methods, we skip most of them here.)

An adaptation of lexicographic goal programming for convex problems is provided in Caballero et al. (1996). The idea is to produce satisfying solutions by solving the hybrid problem (in Section 3.3) with the components of the goal programming solution as upper bounds. Varying the weights produces different solutions.

Lexicographic goal programming is modified significantly in Caballero et al. (1997). No deviational variables are used and the objective function of each priority level is optimized at each iteration. The approach is valid for convex problems.

A solution method for lexicographic goal programming problems where objective functions are fractions of linear or nonlinear functions is described in Pal and Basu (1995). More than one objective function can then belong to the same priority class. The method has characteristics of dynamic programming.

A generalized reduced gradient (GRG) method-based solution algorithm for lexicographic and weighted nonlinear goal programming problems is introduced in Saber and Ravindran (1996). This partitioning technique is demonstrated to be reliable and robust. Several aspects to take into account when aiming at the efficient implementation of goal programming approaches are collected in Tamiz and Jones (1996).

Goal programming can be expanded in an interactive direction in different ways. One can systematically modify the weighting vectors or the lexicographic order of the objective functions or ask for new aspiration levels from the decision maker. These topics are considered in Tamiz and Jones (1997a, b).

4.3.5. Concluding Remarks

Goal programming is a very widely used and popular solution method for practical multiobjective optimization problems. One of the reasons is its age. Another reason is that goal-setting is an understandable and easy way of making decisions. The specification of the weighting coefficients or the lexicographic ordering may be more difficult. The weights do not have so direct an effect on the solution obtained as in the a priori weighting method. However, they are relative to each other. This means that only the relations of the weighting coefficient matter, not the weights themselves. It may be difficult to specify the weights because they have no direct physical meaning. It is demonstrated in Nakayama (1995) that desirable solutions are very difficult to obtain by adjusting the weighting coefficients in the weighted goal programming problem. Anyway, it is as advisable as in the weighting method to normalize the objective functions when weighting coefficients are used.

One must be careful with the selection of the aspiration levels so that the Pareto optimality of the solutions can be guaranteed. The correct selection may be difficult for a decision maker who does not know what the feasible region looks like. Presenting the ranges of the Pareto optimal set, or at least the ideal objective vector, to the decision maker may help in the selection.

Goal programming is not an appropriate method to use if it is desired to obtain trade-offs. Another restricting property is the underlying assumption of a piecewise linear value function and thus piecewise constant marginal rates of substitution.

Assuming that goal programming follows a traditional product life cycle, it is inferred in Schniederjans (1995b) that the current stage of productivity is in decline. It is pointed out that the number of goal programming papers has been on the decrease for several years. One of the reasons suggested is the aging of the few active contributors to goal programming.

Part III

RELATED ISSUES

5. INTERACTIVE METHODS

The class of interactive methods is the most developed of the four classes of methods presented here. The interest devoted to this class can be explained by the fact that assuming the decision maker has enough time and capabilities for co-operation, interactive methods can be presumed to produce the most satisfactory results. Many of the weak points of the methods in the other three classes are overcome. Namely, only part of the Pareto optimal points has to be generated and evaluated, and the decision maker can specify and correct her or his preferences and selections as the solution process continues and (s)he gets to know the problem and its potentialities better. This also means that the decision maker does not have to know any global preference structure. In addition, the decision maker can be assumed to have more confidence in the final solution since (s)he is involved throughout the solution process.

In interactive methods, the decision maker works together with an analyst or an interactive computer program. One can say that the analyst tries to determine the preference structure of the decision maker in an interactive way. A solution pattern is formed and repeated several times. After every iteration, some information is given to the decision maker and (s)he is asked to answer some questions or provide some other type of information. The working order in these methods is: 1) analyst, 2) decision maker, 3) analyst, 4) decision maker, etc. After a reasonable (finite) number of iterations every interactive method should yield a solution that the decision maker can be satisfied with and convinced that no considerably better solution exists. The basic steps in interactive algorithms can be expressed as

a) find an initial feasible solution,
b) interact with the decision maker, and
c) obtain a new solution (or a set of new solutions). If the new solution (or one of them) or one of the previous solutions is acceptable to the decision maker, stop. Otherwise, go to step b).

Interactive methods differ from each other by the form in which information is given to the decision maker, by the form in which information is provided by the decision maker, and how the problem is transformed into a single objective optimization problem. One problem to be solved when designing an interactive method is what kind of data one should use to interact with the decision maker. It should be meaningful and easy for the decision maker to comprehend. The

K. Miettinen, *Nonlinear Multiobjective Optimization*

decision maker should understand the meaning of the parameters for which (s)he is asked to supply values. On the other hand, the data provided to the decision maker should be easily obtainable by the analyst and contain information about the system. Too much information should not be used and the information obtained from the decision maker should be utilized efficiently. To ensure that the greatest possible benefit can be obtained from the interactive method, the decision maker must find the method worthwhile and acceptable, and (s)he must be able to use it properly. This usually means that the method must be understandable and sufficiently easy to use. This aim calls for research in understanding the underlying decision processes and how decisions are made.

As stressed in Kok (1986), experiments in psychology indicate that the amount of information provided to the decision maker has a crucial role. If more information is given to the decision maker, the percentage of the information used decreases. In other words, more information is not necessarily better than less information. More information may increase the confidence of the decision maker in the solution obtained but the quality of the solution may nonetheless be worse.

In addition to the fact that the decision maker has an essential role in interactive methods, the analyst should not be forgotten either. The analyst can support the decision maker in many ways and, in the best possible case, explain the behaviour of the problem to the decision maker. Thus, the analyst may play a meaningful role in the learning process of the decision maker.

Interactive methods have been classified in many ways, mainly according to their solution approaches. Here we do not follow any of those classifications. Let us, however, mention two different conceptions regarding interactive approaches according to Vanderpooten (1989a, b, 1992). The approaches are searching and learning. In *searching-oriented methods* a converging sequence of solution proposals is presented to the decision maker. It is assumed that the decision maker provides consistent preference information. In *learning-oriented methods* a free exploration of alternatives is possible allowing trial and error. The latter does not guide the decision maker and convergence is not guaranteed. The best procedure would be a combination of these two approaches, drawing on their positive features. Such an approach would support the learning of preferences, while it would also include guiding properties.

Before we present any methods, some critical comments are in order. Repeatedly, it has been and will be assumed that the decision maker makes consistent decisions or that (s)he has an underlying (implicitly known) value function upon which her or his decisions are made. The purpose is not to go deeply into the theories of decision making. However, it is worth mentioning that those assumptions can be called into question because they are difficult to verify.

Consistency of the responses of the decision maker is one of the most important factors guaranteeing the success of many interactive solution methods. Because of the subjectivity of the decision makers, different starting points, different types of questions or interaction styles may lead to different final so-

lutions. Some methods are more sensitive with respect to consistency than others. The handling of inconsistency with respect to several interactive methods is treated in Shin and Ravindran (1991). In general, inconsistency can be reduced by consistency tests during the solution process or by minimizing the decision maker's cognitive burden. In other words, interactive methods assuming consistent answers should have built-in mechanisms to deal with inconsistencies. This is one of the motivations in developing new methods for multiobjective optimization.

Further, once the existence of an underlying, implicit value function is supposed, several assumptions are set on it. How can one guarantee and verify, for example, the pseudoconcavity of a function that is not explicitly known? Naturally, something can be concluded if we find out enough about the decision maker's preference structure. Steps in that direction are, however, very laborious and in any case the results are likely to be controversial.

In solving practical problems, knowledge about decision processes and decision analysis is needed to guarantee fruitful co-operation between the decision maker and the analyst. An understanding of the behaviour of the decision maker is important in both developing and applying interactive methods. This fact has been somewhat underestimated, as emphasized in Korhonen and Wallenius (1996, 1997). Korhonen and Wallenius also handle several behavioural issues related to interactive methods. Among them are the learning process of the decision maker, her or his wish to control the search process, and the permissibility of cyclic behaviour or making errors. Perhaps the behavioural sciences should be taken more widely into account when designing interactive methods. A critique of the assumptions underlying interactive methods is also presented in French (1984). The primary concern is that the assumptions should be supported by empirical research from the behavioural sciences.

One noteworthy aspect is that it is unrealistic to assume that decision makers can provide precise information and inputs. After studying 86 reported applications of decision analysis in the literature, it is concluded in Corner and Corner (1995) that the methods should become more user-friendly and descriptive in dealing with the input of the decision maker. In Wierzbicki (1997a), it is stressed that intuition plays an essential role in decision making. Wierzbicki defines intuitive decisions as "quasiconscious and subconscious information processing, leading to an action, utilizing aggregated experience and training and performed (most probably) by a specialized part of the human mind." To provoke intuitive decision making, analysts should provide information in rich and multidimensional graphic terms and avoid requiring consistency.

Decision making is appositely described in Zeleny (1989) as "searching for harmony in chaos." One can criticize the way decision makers are forced into a priori formulas, patterns or contexts (like wandering around the Pareto optimal set). Instead, the decision maker should be guided through her or his own creative search process since decision making can be regarded as a process of continuous redefinition of the problem.

Decision analysis is not handled here in more detail. The above-mentioned aspects are only a few examples of the issues involved.

One more interesting concept is the *convergence* of an interactive method. One can understand several different features as convergence. On the one hand, it may be said that the method converges into Pareto optimal points if the final solution can be proved to be Pareto optimal. One can also say that the method converges into a satisficing solution, if the final solution is satisficing. Finding the best Pareto optimal compromise solution may be understood as convergence. On the other hand, convergence may mean that the final solution is optimal for an underlying value function. This kind of *mathematical convergence* result necessitates certain assumptions about the underlying value function. In this case, the observations mentioned above are valid. If the method is not based on the assumption on any underlying value function, this conception of convergence cannot always be applied.

To sum up, it is not unequivocal what convergence means and how it should be proved. For this reason, it is difficult to provide convergence results for the different methods under consideration. It can also be claimed that mathematical convergence is neither necessary nor sufficient to indicate the practical validity of a method, as stated, for example, in Stewart (1997). The same idea is also expressed in Gardiner and Vanderpooten (1997) and Zionts (1997a, b). On the grounds of the above-mentioned statements, the mathematical convergence properties have been relegated to a secondary position in what follows.

As far as the structure of the methods is concerned, one can require that interactive procedures should converge well immediately in the few initial iterations. This is concluded, for example, in Korhonen et al. (1990) after experimental tests with interactive methods. Decision makers are not willing to wait for progress for a long time.

Stopping criteria are related to the convergence of interactive methods. There are three main stopping criteria. Either the decision maker gets tired of the solution process, some algorithmic stopping (convergence) rule is fulfilled or the decision maker finds a desirable solution and wants to stop. It is difficult to define precisely when a solution is desirable enough to become a final solution.

The convergence of the method has sometimes been considered to be an important factor when selecting a method. However, as stated in Vanderpooten and Vincke (1989), the solution process should not be stopped because of any convergence test. The only practical stopping criterion is the satisfaction of the decision maker with the solution obtained. This usually means that the decision maker must feel that (s)he has received enough information about the problem to be solved.

The current view is that a solution is a final solution if the decision maker is convinced that it is preferred to all the other Pareto optimal solutions (see Korhonen and Wallenius (1996, 1997)). This means that the decision maker must have sufficient evidence that no significantly better solutions exist (see Gardiner and Vanderpooten (1997)). Gardiner and Vanderpooten have studied the interactive solution processes reported in the literature. They point out

that the median number of iterations has been between three and eight. One can ask whether such rapid convergence is the result of getting tired or whether it is due to some other reason. Possibly the decision makers did not know how to continue the solution process.

An important factor when using interactive solution methods is the selection of the starting point. Particularly for nonconvex problems where the objective functions may have several local optima, the starting point affects greatly the solutions generated. If the starting point is somehow biased, it may anchor the desires and the preferences of the decision maker. It is not desirable that the final solution is affected by the starting point. In general, the starting point should provide a useful basis for the decision maker in exploring the Pareto optimal set. The starting point can, for example, be generated by some of the noninteractive methods.

Nonconvexity is a mathematical aspect. Another aspect related to starting points from the point of view of human judgment and decision making is the above-mentioned anchoring. To be more exact, *anchoring* means that the decision maker fixes her or his thinking on some (possible irrelevant) information, like the starting point, and fails to sufficiently adjust and move away from that anchor. In other words, the decision maker is unable to move far from the starting point. This kind of behavioural perspective on interactive decision making is handled in Buchanan and Corner (1997). On the basis of a number of experiments it is argued that anchoring effects are connected more to directed and structured solution methods than to methods based on free search. Buchanan and Corner conclude that whenever an anchoring bias is possible, it is important that the starting point reflects the initial preferences of the decision maker. The reasoning is that since any starting point is likely to bias the decision maker, it is best to bias her or him in the right direction.

Even though interactive methods can be regarded as most promising solution methods for multiobjective optimization problems, there are still cases where these methods are not practicable regardless of the availability of the decision maker. Such problems include, for instance, many engineering problems that require extensive and expensive calculations (like large-scale finite element approximations). One must, however, remember that computational facilities have developed greatly during the last few years. Thus, the number of problems that cannot be solved by interactive methods has decreased. See Osyczka and Zajac (1990) for a suggestion of handling computationally expensive functions. On the other hand, the large number of objective functions may make interactive methods impractical. In this case, it may be difficult for the decision maker to absorb the information provided and to give consistent answers in order to direct the solution process.

Below, we present several interactive methods. Some of them are relatively old and much tested and developed, whereas some others are new and deserve further refinement. The methods to be described are the interactive surrogate worth trade-off method, the Geoffrion-Dyer-Feinberg method, the sequential

proxy optimization technique, the Tchebycheff method, the step method, the reference point method, the GUESS method, the satisficing trade-off method, the light beam search, the reference direction approach, the reference direction method and the NIMBUS method. The first three methods are based on the existence of an underlying value function, whereas the last eight use reference points and the classification of the objectives. (In developing the last of these, attempts have been made to overcome some of the drawbacks observed in the other methods.)

All the methods to be presented are based on generating mainly weakly, properly or Pareto optimal solutions. In each method, it is assumed that less is preferred to more by the decision maker. The same notion could be formulated to require that the underlying value function is strongly decreasing. The reason for avoiding this wording is that an underlying value function is not always assumed to exist. The assumption only concerns the form of the general preference structure of the decision maker.

In connection with the methods, some applications reported in the literature are mentioned. However, let us keep in mind that the impressions obtained from such applications may be biased because unsuccessful applications are hardly ever published. In addition, we give references for extensions and modifications of the methods. We also indicate whether the methods belong to the class of ad hoc or non ad hoc methods. (These classes were introduced at the beginning of this part in Chapter 1.)

Throughout the book the *iteration counter* is denoted by h and the decision variable vector at the current iteration by $\mathbf{x}^h$. In addition, the number of alternative objective vectors presented to the decision maker is denoted by P.

5.1. Interactive Surrogate Worth Trade-Off Method

The interactive surrogate worth trade-off (ISWT) method, put forward in Chankong and Haimes (1978, 1983b, pp. 371–379), is an extension of the surrogate worth trade-off (SWT) method presented in Haimes and Hall (1974) and Haimes et al. (1975). We do not go into details of the SWT method here, but present directly the interactive version. The motivation for including the ISWT method in this book is that it is of theoretical interest.

5.1.1. Introduction

The ε-constraint method, introduced in Section 3.2, is a fundamental element of the ISWT method. The idea is to maximize an underlying (implicitly known) value function. The opinions of the decision maker concerning the trade-off rates at the current solution point determine a search direction. The step-size to be taken in the search direction is determined by solving several

ε-constraint problems and asking the decision maker to select the most satisfactory solution for the continuation. In what follows, appropriate assumptions are assumed to be valid so that the solutions produced by the ε-constraint method are Pareto optimal (see Section 3.2).

It is assumed that

1. The underlying value function $U\colon \mathbf{R}^k \to \mathbf{R}$ exists and is implicitly known to the decision maker. In addition, U is continuously differentiable and strongly decreasing.
2. The objective and the constraint functions are twice continuously differentiable.
3. The feasible region S is compact (so that a finite solution exists for every feasible ε-constraint problem).
4. The assumptions in Theorem 3.2.13 are satisfied.

5.1.2. ISWT Algorithm

The main features of the ISWT method can be presented cursorily with four steps.

(1) Select the reference function f_ℓ to be minimized and give upper bounds to the other objective functions. Set $h = 1$.
(2) Solve the current ε-constraint problem to get a Pareto optimal solution $\mathbf{x}^h$. Trade-off rate information is obtained from the connected Karush-Kuhn-Tucker multipliers.
(3) Ask the opinions of the decision maker with respect to the trade-off rates at $\mathbf{z}^h$ corresponding to $\mathbf{x}^h$.
(4) If some stopping criterion is satisfied, stop with $\mathbf{x}^h$ as the final solution. Otherwise update the upper bounds of the objective functions with the help of the answers obtained in step (3) and solve several ε-constraint problems (to determine an appropriate step-size). Let the decision maker choose the most preferred alternative. Denote the corresponding decision vector by $\mathbf{x}^{h+1}$ and set $h = h + 1$. Go to step (3).

First, we examine how trade-off rate information is obtained from Karush-Kuhn-Tucker multipliers. As noted in Theorem 3.2.13 of Section 3.2, the Karush-Kuhn-Tucker multipliers represent trade-off rates under the specified assumptions.

Let $\mathbf{x}^h \in S$ be a solution of the ε-constraint problem at the iteration h, where f_ℓ is the function to be minimized and the upper bounds are ε_i^h for $i = 1, \ldots, k$, $i \neq \ell$. We suppose that $\mathbf{x}^h$ satisfies the assumptions specified in Theorem 3.2.13. If the Karush-Kuhn-Tucker multipliers $\lambda_{\ell i}^h$ associated with the constraints $f_i(\mathbf{x}) \leq \varepsilon_i^h$ are strictly positive for all $i = 1, \ldots, k$, $i \neq \ell$, then $\lambda_{\ell i}^h$ represents the partial trade-off rate at $\mathbf{x}^h$ between f_ℓ and f_i. In other words, if the multiplier $\lambda_{\ell i}^h$ corresponding to the constraint involving f_i is positive, this particular constraint is active and binds the optimum.

We know now that to move from $\mathbf{x}^h$ to some other (locally) Pareto optimal solution in the neighbourhood of $\mathbf{x}^h$, the value of the function f_ℓ decreases by $\lambda^h_{\ell i}$ units for every unit of increment in the value of the function f_i (or vice versa), while the values of all the other objective functions remain unaltered. The opinion of the decision maker with regard to this kind of trade-off rate for all $i = 1, \ldots, k$, $i \neq \ell$, is found out by posing the following question.

Let an objective vector $\mathbf{f}(\mathbf{x}^h) = \mathbf{z}^h$ be given. If the value of f_ℓ is decreased by $\lambda^h_{\ell i}$ units, then the value of f_i is increased by one unit (or vice versa) and the other objective values remain unaltered. How desirable do you find this trade-off?

If the situation is not so convenient as presented above, that is, some of the Karush-Kuhn-Tucker multipliers $\lambda^h_{\ell i}$ equal zero, then another type of question is needed. Let us suppose that $\lambda^h_{\ell i} > 0$ for $i \in N^>$ and $\lambda^h_{\ell j} = 0$ for $j \in N^=$, where $N^> \cup N^= = \{i \mid i = 1, \ldots, k, \ i \neq \ell\}$. As noted in Theorem 3.2.13, increasing the value of f_i, where $i \in N^>$ decreases the value of f_ℓ and in addition, the values of all f_j also change, where $j \in N^=$. The question to the decision maker for all $i \in N^>$ is now of the form

Let an objective vector $\mathbf{f}(\mathbf{x}^h) = \mathbf{z}^h$ be given. If the value of f_ℓ is decreased by $\lambda^h_{\ell i}$ units, then the value of f_i is increased by one unit (or vice versa) and the values of f_j for $j \in N^=$ change by $\nabla f_j(\mathbf{x}^h)^T \frac{\partial \mathbf{x}(\boldsymbol{\varepsilon}^h)}{\partial \varepsilon_i}$ units. How desirable do you find these trade-offs?

A problem with the question above is that the values of $\frac{\partial \mathbf{x}(\boldsymbol{\varepsilon}^h)}{\partial \varepsilon_i}$ for $i \in N^>$ are unknown. One of the ways suggested in Chankong and Haimes (1983b) for coping with this is that the values can be approximated by solving the ε-constraint problem with a slightly modified upper bound vector as $\boldsymbol{\varepsilon}^h(i) = (\varepsilon^h_1, \ldots, \varepsilon^h_{\ell-1}, \varepsilon^h_{\ell+1}, \ldots, \varepsilon^h_i + \epsilon, \ldots, \varepsilon^h_k)$, where $\epsilon \neq 0$ is a scalar with a small absolute value. Let the solution of this ε-constraint problem be $\mathbf{x}(\boldsymbol{\varepsilon}^h(i))$. We obtain now an approximation by

$$\frac{\partial \mathbf{x}(\boldsymbol{\varepsilon}^h)}{\partial \varepsilon_i} \approx \frac{\mathbf{x}(\boldsymbol{\varepsilon}^h(i)) - \mathbf{x}^h}{\epsilon}.$$

Note that the decision maker's opinions are asked respecting certain amounts of change in the values of the objective functions, and not of changes in general. The following problem to be handled is the form of the answers expected from the decision maker. It is suggested in Chankong and Haimes (1978, 1983b) that the decision maker must specify an integer between 10 and -10 to indicate her or his degree of preference. If the decision maker is completely satisfied with the trade-off suggested, the answer is 10. Positive numbers less that 10 indicate the degree of satisfaction (less than complete). Correspondingly, negative answers reflect the decision maker's satisfaction with the trade-off which is converse to that in the question. The answer 0 means that the decision maker is indifferent to the given trade-off.

In Tarvainen (1984), it is suggested that far fewer choices are given to the decision maker. The possible answers are integers from 2 to -2 and their meaning corresponds to that presented above. The justification is that it is easier for the decision maker to give an answer and maintain some kind of consistency when there are fewer alternatives. These five alternatives are enough to represent the direction and rough degree of the decision maker's preferences and satisfaction.

Regardless of the scale selected, the response of the decision maker is called a *surrogate worth* of the trade-off rate between f_ℓ and f_i at $\mathbf{x}^h$ and denoted by $W_{\ell i}^h$. At each point $\mathbf{x}^h$, a number of $k-1$ (or less, if $N^= \neq \emptyset$) questions of the previously described form are presented to the decision maker and the values for $W_{\ell i}^h$ $(i = 1, \ldots, k,\ i \neq \ell)$ are obtained.

According to Theorem 3.2.13, there exists a Pareto optimal solution in the neighbourhood of $\mathbf{x}^h$ when the values of the objective functions are changed according to the information given in the trade-off rates. The problem is how much the values of the objective functions can be changed in order to remain on the Pareto optimal surface and obtain the best possible solution. We must find a way to update the upper bounds of the objective functions in an appropriate way.

How to proceed from this point depends on the scale chosen for the surrogate worth values. The idea is to obtain an estimate for the gradient of the underlying value function with the help of the surrogate worth values. Then a steepest ascent-type formulation is used to maximize the value function. The upper bounds of the ε-constraint problem are revised and a new solution is obtained. It is assumed to satisfy the preferences of the decision maker indicated by the surrogate worth values as well as possible.

In the original version by Chankong and Haimes, it is suggested that the upper bounds are updated from iteration h to $h+1$ by

$$\varepsilon_i^{h+1} = \varepsilon_i^h + t(W_{\ell i}^h | f_i(\mathbf{x}^h)|)$$

for $i \in N^>$ and

$$\varepsilon_j^{h+1} = \varepsilon_j^h + t\left(\nabla f_j(\mathbf{x}^h)^T \left(\frac{\partial \mathbf{x}(\boldsymbol{\varepsilon}^h)}{\partial \varepsilon_i}\right) W_{\ell i}^h | f_i(\mathbf{x}^h)|\right)$$

for $j \in N^=$, where $i \in N^>$ and t is a step-size to be determined. For details, see Chankong and Haimes (1978) and references therein.

For simplicity, it is assumed in Tarvainen (1984) that the Karush-Kuhn-Tucker multipliers are all strictly positive. The decision maker is asked to specify small and meaningful amounts Δf_i for all $i = 1, \ldots, k,\ i \neq \ell$. The scalar Δf_i represents the amount of change in the value of f_i that is relevant to the decision maker. The upper bounds are now updated by

$$\varepsilon_i^{h+1} = \varepsilon_i^h + t(W_{\ell i}^h \Delta f_i)$$

for $i = 1, \ldots, k,\ i \neq \ell$, where t denotes the step-size.

Several discrete values may be given to the step-size t in each updating formula. Then the ε-constraint problem is solved for every value. The resulting objective vectors are presented to the decision maker, who is asked to choose the most preferred one. A graphical representation of the alternatives may be helpful. This topic is handled in Chapter 3 of Part III. After choosing the new solution (and thus an appropriate step-size), trade-off rate information at that solution is obtained from the corresponding Karush-Kuhn-Tucker multipliers (as earlier). The procedure continues by asking the decision maker for the surrogate worth values.

5.1.3. Comments

In practice, when the decision maker is asked to express her or his preferences concerning the trade-off rates, (s)he is implicitly asked to compare the trade-off rates with her or his marginal rates of substitution. (Naturally, the decision maker does not have to be able to specify the marginal rates of substitution explicitly.) If $m_{\ell i} < \lambda_{\ell i}$, then the surrogate worth value is positive (and the contrary respectively). If $m_{\ell i} = \lambda_{\ell i}$ for all $i = 1, \ldots, k$, $i \neq \ell$, meaning $W_{\ell i} = 0$, then the stopping criterion (2.8.1) introduced in Subsection 2.8.2 of Part I is valid. Thus, the condition $W_{\ell i}^h = 0$ for all $i \neq \ell$ is a common stopping criterion for the algorithm. Another possible stopping situation is that the decision maker wants to proceed, but only in an infeasible direction. The latter condition is more difficult to check.

The ISWT method can be classified as non ad hoc in nature. If the value function is known, then the trade-off rates are easy to compare with the marginal rates of substitution. Further, when comparing alternatives, it is easy to select the one with the highest value function value.

The convergence rate of the ISWT method greatly depends on the accuracy and the consistency of the answers of the decision maker. It was pointed out in Section 2.8 of Part I that it is important to select the reference function carefully. This comment is also valid when considering the convergence properties. If there is a sharp limit in the values of the reference function where there is a change in satisfaction from 'very satisfactory' to 'very unsatisfactory,' the solution procedure may stop too early. Further references are cited in Chankong and Haimes (1978) for convergence results.

A method related to the ISWT method is presented in Chen and Wang (1984). The method is an interactive version of the SWT method, where new solution alternatives are generated by Lin's proper equality method (see Section 3.2), and the decision maker has to specify only the sign of the surrogate worth values.

There are many other modifications of the SWT method in the literature. Among others, it is generalized for multiple decision makers in Chankong and Haimes (1983b, pp. 359–366), Haimes (1980) and Hall and Haimes (1976). The first two handle also the SWT method in stochastic problems.

5.1.4. Concluding Remarks

The role of the decision maker is quite easy to understand in the ISWT method. (S)he is provided with one solution and has to specify the surrogate worth values. The complicatedness of giving the answers depends on how experienced the decision maker is in such specification and which variation of the method is employed. The set of 21 different alternatives as surrogate worth values in the original version is quite a lot to select from. It may be difficult for the decision maker to provide consistent answers throughout the decision process. In addition, if there is a large number of objective functions, the decision maker has to specify a lot of surrogate worth values at each iteration. At least for some decision makers it may be easier to maintain consistency when there are fewer alternative values for the surrogate worth available (as suggested by Tarvainen (1984)).

Trade-off rates play an important role in the ISWT method, and that is why the decision maker has to understand the concept of trade-off properly. Attention must also be paid to the ease of understanding and careful formulation of the questions concerning the trade-off rates. Careless formulation may, for example, cause the sign of the surrogate worth value to be changed.

It is a virtue that all the alternatives during the solution process are Pareto optimal. Thus, the decision maker is not bothered with any other kind of solutions.

A negative feature is that there are a lot of different assumptions to be satisfied to guarantee that the algorithm works. It may be difficult (and at least laborious) in many practical problems to ensure that the assumptions are satisfied. One can argue that the validity of the assumptions is not always that important in practice. However, for example, the correctness of the trade-off rates is crucial for the success of the ISWT method.

5.2. Geoffrion-Dyer-Feinberg Method

The Geoffrion-Dyer-Feinberg (GDF) method, proposed in Geoffrion et al. (1972), is an interactive method based in principle on the same idea as the ISWT method; maximization of the underlying (implicitly known) value function. The realization is quite different, though. The GDF method is one of the most well-known interactive methods.

5.2.1. Introduction

The basic idea behind the GDF and the ISWT methods is the same. At each iteration, a local approximation of an underlying value function is generated and maximized. In the GDF method, the idea is somewhat more clearly visible. Marginal rates of substitution specified by the decision maker are used

to approximate the direction of steepest ascent of the value function. Then the value function is maximized by a gradient-based method. A gradient method of Frank and Wolfe (FW) (see Frank and Wolfe (1956)) has been selected for optimization because of its simplicity and robust convergence (rapid initial convergence) properties. The GDF method is also sometimes called an interactive Frank-Wolfe method, because it has been constructed on the basis of the FW method.

The problem to be solved here is

$$\begin{aligned} &\text{maximize} && u(\mathbf{x}) = U(\mathbf{f}(\mathbf{x})) \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{5.2.1}$$

It is assumed that

1. The underlying value function $U\colon \mathbf{R}^k \to \mathbf{R}$ exists and is implicitly known to the decision maker. In addition, $u\colon \mathbf{R}^n \to \mathbf{R}$ is a continuously differentiable and concave function on S (sufficient conditions for the concavity are, for example, that U is a concave decreasing function and the objective functions are convex; or U is concave and the objective functions are linear), and U is strongly decreasing with respect to the reference function (denoted here by f_ℓ) so that $\frac{\partial U(\mathbf{f}(\mathbf{x}))}{\partial f_\ell} < 0$.
2. The objective functions are continuously differentiable.
3. The feasible region S is compact and convex.

Let us begin by presenting the main principles of the FW method. Let a point $\mathbf{x}^h \in S$ be given. The idea of the FW method is that when maximizing some objective function $u\colon \mathbf{R}^n \to \mathbf{R}$ subject to constraints $\mathbf{x} \in S$, instead of u, a linear approximation of it at some point $\mathbf{x}^h \in S$ is optimized. If the solution obtained is $\mathbf{y}^h$, then the direction $\mathbf{d}^h = \mathbf{y}^h - \mathbf{x}^h$ is a promising direction in which to seek an increased value for the objective function u.

At any feasible point $\mathbf{x}^h$, a linear approximation to $u(\mathbf{y})$ is

$$u(\mathbf{x}^h) + \nabla_x u(\mathbf{x}^h)^T(\mathbf{y} - \mathbf{x}^h).$$

The maximization of this linear approximation, after excluding constant terms, is equivalent to the problem

$$\begin{aligned} &\text{maximize} && \nabla_x u(\mathbf{x}^h)^T\mathbf{y} \\ &\text{subject to} && \mathbf{y} \in S, \end{aligned} \tag{5.2.2}$$

where $\mathbf{x}^h$ is fixed and $\mathbf{y} \in \mathbf{R}^n$ is the variable. Let $\mathbf{y}^h \in S$ be the solution.

A well-known condition for $\mathbf{x}^h$ to be an optimal solution of problem (5.2.1) is that $\nabla_x u(\mathbf{x}^h)^T\mathbf{d} \leq 0$ for all $\mathbf{d} \in S$. Therefore, if after solving problem (5.2.2) is $\mathbf{y}^h = \mathbf{x}^h$, then we know that $0 = \nabla_x u(\mathbf{x}^h)^T(\mathbf{y}^h - \mathbf{x}^h) \geq \nabla_x u(\mathbf{x}^h)^T(\mathbf{y} - \mathbf{x}^h)$ for all $\mathbf{y} \in S$, and, thus, the optimality condition is fulfilled at $\mathbf{x}^h$.

If $\mathbf{y}^h \neq \mathbf{x}^h$, then we set $\mathbf{d}^h = \mathbf{y}^h - \mathbf{x}^h$. The points $\mathbf{y}^h$ and $\mathbf{x}^h$ are feasible and, because of the convexity assumption of S, any new point $\mathbf{x}^{h+1} = \mathbf{x}^h + t\mathbf{d}^h$

where $0 \leq t \leq 1$ is feasible. Finally, we must determine an appropriate step-size in the direction $\mathbf{d}^h$ by maximizing $u(\mathbf{x}^h + t\mathbf{d}^h)$ subject to $0 \leq t \leq 1$.

5.2.2. GDF Algorithm

Below, we shall show that even though we do not know the value function explicitly, we can obtain a local linear approximation for it or to be more exact, its gradient, with the help of marginal rates of substitution. This is enough to permit the FW method to be applied. Before going into details we present the basic phases of the GDF algorithm.

(1) Ask the decision maker to specify a reference function f_ℓ. Choose a feasible starting point $\mathbf{x}^1$. Set $h = 1$.
(2) Ask the decision maker to specify marginal rates of substitution between f_ℓ and the other objectives at the current solution point $\mathbf{x}^h$.
(3) Solve problem (5.2.3), where the approximation of the value function is maximized. Denote the solution by $\mathbf{y}^h \in S$. Set the direction $\mathbf{d}^h = \mathbf{y}^h - \mathbf{x}^h$. If $\mathbf{d}^h = \mathbf{0}$, go to step (6).
(4) Determine with the help of the decision maker the appropriate step-size t^h to be taken in the direction $\mathbf{d}^h$. Denote the corresponding solution by $\mathbf{x}^{h+1} = \mathbf{x}^h + t^h\mathbf{d}^h$.
(5) Set $h = h + 1$. If the decision maker wants to continue, go to step (2).
(6) Stop. The final solution is $\mathbf{x}^h$.

In the algorithm above we need a local linear approximation of the value function at the point $\mathbf{x}^h$. As explained earlier, we only need to know the gradient of the value function at $\mathbf{x}^h$. According to the chain rule, we know that the gradient of the objective function of problem (5.2.1) at the point $\mathbf{x}^h \in S$ can be written in the form

$$\nabla_x U(\mathbf{f}(\mathbf{x}^h)) = \sum_{i=1}^{k} \left(\frac{\partial U(\mathbf{f}(\mathbf{x}^h))}{\partial f_i} \right) \nabla_x f_i(\mathbf{x}^h).$$

In assumption 1 we supposed that $\frac{\partial U(\mathbf{f}(\mathbf{x}^h))}{\partial f_\ell} < 0$, where f_ℓ is the reference function. Positive scaling does not affect the direction of the gradient, so we can divide the gradient of the value function by a positive scalar $-\frac{\partial U(\mathbf{f}(\mathbf{x}^h))}{\partial f_\ell}$. We have now the direction of the gradient of the value function at the point $\mathbf{x}^h$ in the form

$$\sum_{i=1}^{k} -m_i^h \nabla_x f_i(\mathbf{x}^h),$$

where $m_i^h = \frac{\partial U(\mathbf{f}(\mathbf{x}^h))}{\partial f_i} / \frac{\partial U(\mathbf{f}(\mathbf{x}^h))}{\partial f_\ell}$ for all $i = 1, \ldots, k$, $i \neq \ell$. The numbers m_i^h $(= m_{\ell i})$ represent the marginal rates of substitution at $\mathbf{x}^h$ between f_ℓ and f_i (see Remark 2.8.7 of Part I).

The role of the reference function is significant, because marginal rates of substitution are generated with respect to it. The decision maker must be asked to specify the reference function so that the marginal rates of substitution are sensible. Note that if the underlying value function is linear, then only one iteration is needed to achieve the final solution (the marginal rates of substitution are constant).

It may be difficult for the decision maker to specify the marginal rates of substitution directly (or straight away). If this is the case, some auxiliary procedures may be brought in to assist. One such procedure is presented in Dyer (1973a). The idea there is to determine (at the point $\mathbf{f}(\mathbf{x}^h)$) small amounts of f_ℓ and f_i, denoted by Δ_{f_ℓ} and Δ_{f_i}, respectively, such that an increase in the value of f_i by Δ_{f_i} is matched for the decision maker by a compensatory decrease by Δ_{f_ℓ} in the value of f_ℓ, while the values of all the other objective functions remain unaltered. In other words, the vectors $(f_1(\mathbf{x}^h), \ldots, f_k(\mathbf{x}^h))^T$ and $(f_1(\mathbf{x}^h), \ldots, f_\ell(\mathbf{x}^h) - \Delta_{f_\ell}, \ldots, f_i(\mathbf{x}^h) + \Delta_{f_i}, \ldots, f_k(\mathbf{x}^h))^T$ are indifferent to the decision maker. We obtain now

$$m_i^h \approx \frac{\Delta_{f_\ell}}{\Delta_{f_i}},$$

where the approximation becomes arbitrarily exact when the Δ-amounts of change approach 0. Note that $m_\ell^h = 1$.

The approximation of marginal rates of substitution is illustrated in Figure 5.2.1. The bold curve is a contour of the value function and the continuous line its tangent at $\mathbf{z}^h$. The marginal rate of substitution at $\mathbf{z}^h$ is the negative of the slope of that tangent. The slope of the approximating broken line is quite different.

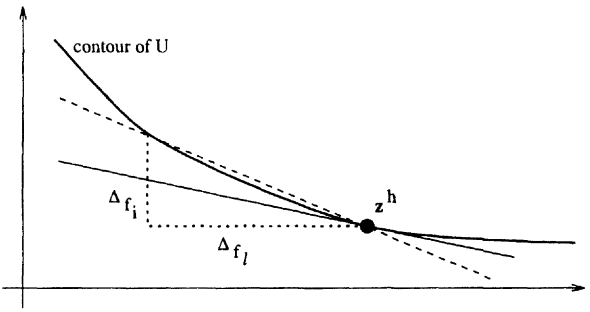

Figure 5.2.1. An approximation of the marginal rate of substitution.

In practice, the Δ-amounts of change cannot be made arbitrarily small near 0, as emphasized in Sawaragi et al. (1985, pp. 259–260). The reason is that human beings cannot recognize small changes beyond a certain point. This threshold of human recognition is called a *just noticeable difference*. That is why the marginal rates of substitution are always approximations of the correct

values. An example of the effects of the just noticeable difference is given in Nakayama (1985a) by illustrating how the solution process may terminate at a wrong solution. For this reason one may have doubts about the adequacy of marginal rates of substitution as a means of providing preference information. They seem to be difficult for the decision maker to specify and their accuracy is questionable.

However, we must now assume that the marginal rates of substitution are provided accurately enough. According to the FW method, the maximization of the linear approximation of U is equivalent to the problem

$$\begin{aligned} &\text{maximize} && \left(\sum_{i=1}^{k} -m_i^h \nabla_x f_i(\mathbf{x}^h)\right)^T \mathbf{y} \\ &\text{subject to} && \mathbf{y} \in S \end{aligned} \tag{5.2.3}$$

with $\mathbf{y} \in \mathbf{R}^n$ being the variable. The solution is denoted by $\mathbf{y}^h$. The existence of the optimal solution is ensured by the compactness of S and the continuity of all the functions.

The search direction is now $\mathbf{d}^h = \mathbf{y}^h - \mathbf{x}^h$. Provided that the marginal rates of substitution are reasonably accurate, the search direction should be usable. Let us mention that a scaling idea presented in Clinton and Troutt (1988) can be included in the method. Heterogeneous objective functions can be scaled to have equal effect in problem (5.2.3) by adjusting the norms of the gradients of the objective functions with scalar coefficients.

The following problem is to find an appropriate step-size for going in the search direction. The only variable is the step-size. The decision maker can be offered objective vectors, where $z_i = f_i(\mathbf{x}^h + t\mathbf{d}^h)$ for $i = 1, \ldots, k$, and t varies stepwise between 0 and 1 (e.g., $t = \frac{j-1}{P-1}$ where $j = 1, \ldots, P$, and P is the number of the alternative objective vectors to be presented). Another possibility is to draw the objective values as a function of t, provided no serious scaling problems exist. An example of the graphical presentation is given in Hwang and Masud (1979, p. 109). Graphical illustration of the alternative objective vectors is handled in Chapter 3 of Part III. Note that the alternatives are not necessarily Pareto optimal. From the information given to the decision maker (s)he selects the most preferred objective vector and the corresponding value of t is selected as t^h. It is obvious that the task of selection becomes more difficult for the decision maker as the number of objective functions increases.

The opinions of the decision maker and the situation $\mathbf{y}^h = \mathbf{x}^h$ are used here as stopping criteria. Other possible criteria are presented in Hwang and Masud (1979, pp. 108–110) and Yu (1985, p. 327).

5.2.3. Comments

The GDF method can be characterized to be a non ad hoc method. If one knows the value function, it is easy to specify the marginal rates of substitution and select the best alternative. The mathematical convergence properties of the GDF method are closely related to the convergence properties of the FW method. The convergence of the FW algorithm under the assumptions provided at the beginning of this section, is proved in Zangwill (1969). However, it must be kept in mind that the correctness of the marginal rates of substitution and the step-sizes affects the convergence considerably. If it is assumed that the answers of the decision maker become ever more exact as the solution process continues, it is asserted in Geoffrion et al. (1972) that infinite convergence holds.

More important than infinite convergence in an interactive procedure like this is the initial rate of convergence, since a satisfactory solution should be found in a reasonable number of iterations. It is claimed in Geoffrion et al. (1972) that the error in the objective function values is at least halved at each of the first H iterations (H is unknown). The convergence becomes slower near the optimum because of the zig-zag-phenomenon.

The effects of errors in estimating the gradient of the value function are investigated in Dyer (1974). The result is that even if the answers of the decision maker are not strictly consistent and the just noticeable difference affects the marginal rates of substitution, the method is stable and converges (only slower) under certain assumptions.

5.2.4. Applications and Extensions

The GDF method is applied in Geoffrion et al. (1972) to the operation of an academic department. Numerical examples are also given, for example, in Hwang and Masud (1979, pp. 111–121) and Steuer (1986, pp. 377–379). A time-sharing computer program implementing the GDF algorithm is suggested in Dyer (1973a). The GDF method is implemented for convex problems by a so-called projection-relaxation procedure in the objective space in Ferreira and Machado (1996). An application in water resources allocation is also given. The GDF method is adapted for continuous equilibrium network design problems in Friesz (1981).

In Dyer (1972), a method called interactive goal programming is presented. It is a combination of the GDF method and goal programming. The vector $\mathbf{y}^h$ is obtained by the means of weighted goal programming with the marginal rates of substitution as weights. Some convergence results are also given. The GDF method and the interactive goal programming method are applied in Jedrzejowicz and Rosicka (1983) to multiobjective reliability optimization problems appearing in multiple classes of system failures.

The GDF method has been a subject of many modifications in the literature. New versions have been mainly developed to overcome some of the

weaknesses of the GDF method. In Hemming (1981), a simplex-based direction-finding problem is proposed for MOLP problems to avoid the specification of the marginal rates of substitution. It is stressed that the convergence properties may be impaired, but the cognitive burden placed on the decision maker is diminished. A revised step-size problem is also presented to produce Pareto optimal solutions. In addition, the GDF method is modified for MOLP problems in Winkels and Meika (1984) so that when determining the step-size at each iteration, the objective vectors are projected with a so-called efficiency projection onto the Pareto optimal set. This is done by solving a parametric linear programming problem.

The GDF method is altered in Rosinger (1981) by constructing a wide family of possible inquiry patterns to lead into the determination of the marginal rates of substitution. The decision maker can choose the form of the inquiry at each iteration. The convergence of the method is also proved.

The so-called proxy approach is introduced in Oppenheimer (1978). The value function is no longer approximated linearly. The idea is to give a local proxy to the value function at each iteration. A sum-of-powers or a sum-of-exponentials proxy is fitted locally by specifying the parameters connected to the problem. Now, direction finding and step-size determination problems are replaced by the maximization of the proxy function. The proxy is not a valid approximation globally, but when used locally, it gives a higher convergence rate than the original GDF method. Even this method does not guarantee the Pareto optimality of the solutions. Oppenheimer does not establish any systematic procedure for maximizing the proxy function. A method improving on and utilizing Oppenheimer's ideas is presented in Section 5.3.

Several modifications of the GDF method are presented in Sadagopan and Ravindran (1986). First, the FW method is replaced by a generalized reduced gradient method. Then, the role of the decision maker is facilitated by asking for intervals for the marginal rates of substitution instead of exact values. The step-size is computed with the help of upper and lower bounds for the objective functions without the decision maker.

In Musselman and Talavage (1980), the idea of the adaptation is to reduce the feasible region according to the marginal rates of substitution specified by the decision maker. Solutions with lower values of the value function than the current solution are dropped. The method permits sensitivity analysis of the decision maker's inputs.

Ideas of the GDF method are applied in the interactive integrated approach for quasiconcave value functions in Al-alvani et al. (1992). A large set of Pareto optimal solutions is first generated and the form of the underlying implicit value function is deduced with pairwise comparisons. The gradient of the value function is also obtained from the comparisons and, thus, marginal rates of substitution are not needed. Solutions along the search direction are projected by the weighted Tchebycheff metric. The stopping criterion is based on trade-off information.

Finally, we mention a modification of the GDF method, known as the subgradient GDF method, for nondifferentiable multiobjective optimization problems, presented in Miettinen (1994) and Miettinen and Mäkelä (1991, 1993, 1994). The twice continuous differentiability of the objective functions is relaxed and they are assumed to be locally Lipschitzian, but the value function has to still be continuously differentiable. The FW method is replaced by the subgradient method (see Shor (1985)) in optimizing the approximated value function.

In addition to being able to handle nondifferentiable functions, the modification has another advantage. It produces only Pareto optimal solutions, unlike the original GDF method. Each calculated solution is set as a reference point to an order-approximating achievement scalarizing function. We know from Section 3.5 that the solutions of such achievement functions are always Pareto optimal. Naturally, additional optimizations increase the computational burden but this is the price to be paid for the certainty that the decision maker does not have to handle non-Pareto optimal solutions. (A strongly decreasing value function implies that less is preferred to more in the mind of the decision maker.)

Some applications solved with the subgradient GDF method are presented in Miettinen (1994). The subgradient GDF method is used in solving an optimal control problem concerning an elastic string in Miettinen and Mäkelä (1993) and continuous casting of steel in Miettinen and Mäkelä (1994).

5.2.5. Concluding Remarks

In the GDF method the decision maker is first given one solution where (s)he has to specify the marginal rates of substitution. After that the decision maker must select the most preferred solution from a set of alternatives. Thus, the ways of interaction are versatile.

In spite of the plausible theoretical foundation of the GDF method, it is not so convincing and powerful in practice. The most important difficulty for the decision maker is the determining of the $k - 1$ marginal rates of substitution at each iteration. Even more difficult is to give consistent and correct marginal rates of substitution at every iteration. The difficulties of the decision maker in determining the marginal rates of substitution are demonstrated, for example, in Wallenius (1975) by comparative tests. The same point can be illustrated by an example from Hemming (1981) where a politician is asked to specify the exact marginal rate of substitution between unemployment and a decrease of 1 % in the inflation rate.

A drawback of the GDF method is that the final solution obtained is not necessarily Pareto optimal. Naturally, it can always be projected onto the Pareto optimal set with an auxiliary problem. A more serious objection is that when several alternatives are given to the decision maker from which to select the step-size, it is likely that many of them are not Pareto optimal. They can also be projected onto the Pareto optimal set before presentation to the decision

maker, but this necessitates extra effort. The projection may be done, for instance, by lexicographic ordering or by the means presented in Section 2.10 of Part I. The use of achievement functions is demonstrated in the subgradient GDF method. The weakness in the projection is that the computational burden increases. It is for the analyst and the decision maker to decide which of the two shortcomings is less inconvenient.

Theoretically, the Pareto optimality of the final solution is guaranteed if the value function is strongly decreasing (by Theorem 2.6.2 of Part I). In any case, marginal rates of substitution are crucial in approximating the value function, and for many decision makers they are difficult and troublesome to specify.

For many people it is easier to think of desired changes in the objective function values than to specify indifference relations. This may, especially, be the case if the objective vector at which the marginal rates of substitution are to be specified is not particularly desirable. Then it may be frustrating to think of indifferent solutions instead of the improvements sought.

The Frank-Wolfe gradient method has been selected as the maximization algorithm for its fast initial convergence. In some cases, other gradient-based methods may be more appropriate. For example, the subgradient method is employed in the subgradient GDF method.

There are a lot of assumptions that the problem to be solved must satisfy in order the method to work and converge. Several sufficient conditions on the decision maker's preferences are presented in Sawaragi et al. (1985, pp. 258–259) to guarantee the differentiability and the concavity of the value function. Even these conditions are not very easy to check. For more critical discussion concerning the GDF method, see Sawaragi et al. (1985, pp.257–261).

5.3. Sequential Proxy Optimization Technique

Like the two previous methods, the sequential proxy optimization technique (SPOT), presented in Sakawa (1982), is based on the idea of maximizing the decision maker's underlying value function, which is once again assumed to be known implicitly. SPOT includes some properties of the ISWT and the GDF methods, and that is why we describe it here briefly.

5.3.1. Introduction

As in the two interactive methods presented thus far, the search direction in SPOT is obtained by approximating locally the gradient of the underlying value function, and the step-size is determined according to the preferences of the decision maker. Here, both marginal rates of substitution and trade-off rates are used in approximating the value function.

It is assumed that

1. The underlying value function $U\colon \mathbf{R}^k \to \mathbf{R}$ exists and is implicitly known to the decision maker. In addition, U is a continuously differentiable, strongly decreasing and concave function on the subset of Z where the points are Pareto optimal.
2. The objective and the constraint functions are convex and twice continuously differentiable.
3. The feasible region S is compact and convex (and there exist some upper bounds for the ε-constraint problem so that the solution is finite).
4. The assumptions in Theorem 3.2.13 are satisfied.

The ε-constraint problem is used to generate Pareto optimal solutions. The solution of ε-constraint problem (3.2.1) is denoted by $\mathbf{x}^h$. It is assumed to be unique so that Pareto optimality is guaranteed. Throughout this section it is assumed that all the upper bound constraints are active at the optimum. (If this is not the case, then the upper bounds must be slightly modified.) Then, $f_j(\mathbf{x}^h) = \varepsilon_j^h$ for all $j = 1, \ldots, k$, $j \neq \ell$. The optimal value of f_ℓ, that is, $f_\ell(\mathbf{x}^h)$, is denoted by z_ℓ^h. It is also assumed that all the Karush-Kuhn-Tucker multipliers associated with the active constraints are strictly positive. The conditions of Theorem 3.2.13 are assumed to be satisfied so that trade-off rate information can be obtained from the Karush-Kuhn-Tucker multipliers.

Here, the value function is not maximized in form (4.1.1) as before. Instead, the set of feasible alternatives is restricted to the Pareto optimal set. According to the assumption above stating that $f_j(\mathbf{x}^h) = \varepsilon_j^h$ for all $j = 1, \ldots, k$, $j \neq \ell$, we have a new formulation:

$$\text{(5.3.1)} \qquad \text{maximize} \quad U(\varepsilon_1^h, \ldots, \varepsilon_{\ell-1}^h, z_\ell^h, \varepsilon_{\ell+1}^h, \ldots, \varepsilon_k^h).$$

No constraints are needed here since the formulation includes the original constraints. The optimization is now carried out in the objective space $\mathbf{R}^{k-1}$, where the upper bounds ε_j^h are the variables.

It is proved in Sakawa (1982) that the new function is concave with respect to those $\boldsymbol{\varepsilon} \in \mathbf{R}^{k-1}$ for which the upper bound constraints are all active. Sakawa also claims that the partial derivative of (5.3.1) with respect to ε_j^h, $j = 1, \ldots, k$, $j \neq \ell$, is equivalent to $\frac{\partial U(\cdot)}{\partial f_\ell}(m_{\ell j}^h - \lambda_{\ell j}^h)$, where $m_{\ell j}^h$ is the marginal rate of substitution between f_ℓ and f_j at $\mathbf{x}^h$ (obtained from the decision maker, see Section 5.2) and $\lambda_{\ell j}^h$ is the partial trade-off rate between f_ℓ and f_j at $\mathbf{x}^h$ (obtained from the Karush-Kuhn-Tucker multipliers, see Sections 3.2 and 5.1).

Because it was assumed that the value function is strongly decreasing, we know that $\frac{\partial U(\cdot)}{\partial f_\ell} < 0$ and we can divide by it. We denote now

$$\Delta\varepsilon_j^h = -(m_{\ell j}^h - \lambda_{\ell j}^h)$$

for $j = 1, \ldots, k$, $j \neq \ell$, and it represents the direction of steepest ascent of the value function (5.3.1) at the current point $\mathbf{x}^h$ for $j \neq \ell$. According to Sakawa, the ℓth component of the direction is

$$\sum_{\substack{j=1\\ j\neq\ell}}^{k} \lambda_{\ell j}^{h}(m_{\ell j}^{h} - \lambda_{\ell j}^{h}) = \sum_{\substack{j=1\\ j\neq\ell}}^{k} -\lambda_{\ell j}^{h}\Delta\varepsilon_{j}^{h}$$

denoted by Δz_{ℓ}^{h}.

After obtaining the search direction, we have to find the step-size t which in theory maximizes the function

$$\begin{aligned} U(&\varepsilon_1^h + t\Delta\varepsilon_1^h, \ldots, \varepsilon_{\ell-1}^h + t\Delta\varepsilon_{\ell-1}^h, z_\ell^h + t\Delta z_\ell^h, \\ &\varepsilon_{\ell+1}^h + t\Delta\varepsilon_{\ell+1}^h, \ldots, \varepsilon_k^h + t\Delta\varepsilon_k^h). \end{aligned} \tag{5.3.2}$$

The step-size could be determined as earlier by presenting different objective vectors to the decision maker with different values of t and by letting the decision maker choose the most preferred one. The problem with alternative objective vectors of this kind is that they are not necessarily Pareto optimal.

However, the step-size is not determined by asking the decision maker's opinion. Nonetheless, different alternatives are generated with different step-sizes. Their Pareto optimality is guaranteed by solving the ε-constraint problem with the upper bounds $\varepsilon_j^h + t\Delta\varepsilon_j^h$ for $j = 1, \ldots, k$, $j \neq \ell$, still assuming that the constraints are active. This increases the number of calculations since the ε-constraint problem must be solved for several values of t.

The best alternative is selected employing local proxy preference functions p (in the same spirit as the proxy approach presented in Section 5.2 in connection with the GDF method). The proxy function replaces function (5.3.2) and the alternative with the highest proxy function value is selected for the continuation.

According to the preference structure of the decision maker a sum-of-exponentials, sum-of-powers or sum-of-logarithms proxy function of the form

$$-\sum_{i=1}^{k} a_i e^{-\omega_i f_i(\mathbf{x})}, \quad -\sum_{i=1}^{k} a_i (n_i + f_i(\mathbf{x}))^{\alpha_i} \quad \text{or} \quad \sum_{i=1}^{k} a_i \ln(n_i - f_i(\mathbf{x})),$$

respectively, is used. The constants a_i, ω_i, n_i and α_i are used to tune the proxy functions so that they represent the current problem and the preferences of the decision maker better, and they are derived from the marginal rates of substitution; see, for example, Sakawa (1982) and Sakawa and Seo (1982b) for further details. This kind of proxy function is very restrictive globally but reasonable when assumed locally.

5.3.2. SPOT Algorithm

We can now present the basic ideas of the SPOT algorithm.

(1) Choose a reference function f_ℓ and upper bounds $\boldsymbol{\varepsilon}^1 \in \mathbf{R}^{k-1}$ for which all the constraints of the ε-constraint problem are active. Set $h = 1$.

(2) Solve the current (active) ε-constraint problem for $\boldsymbol{\varepsilon}^h$ to obtain a solution $\mathbf{x}^h$.

(3) Denote the Pareto optimal objective vector corresponding to $\mathbf{x}^h$ by $\mathbf{z}^h$ and the corresponding Karush-Kuhn-Tucker multipliers by $\lambda_{\ell j}^h$, $j = 1, \ldots, k$, $j \neq \ell$.

(4) Ask the decision maker for the marginal rates of substitution $m_{\ell j}^h$ for $j = 1, \ldots, k$, $j \neq \ell$, at $\mathbf{x}^h$. Test the consistency of the marginal rates of substitution and ask the decision maker to respecify them if necessary.

(5) If $|m_{\ell j}^h - \lambda_{\ell j}^h| < \theta$, where θ is a prespecified positive tolerance, then stop with $\mathbf{x}^h$ as the final solution. Otherwise, determine the components $\Delta\varepsilon_j^h$, $j \neq \ell$, of the search direction vector.

(6) Select an appropriate form of the proxy function and calculate its parameters. If the obtained proxy function is not strongly decreasing and concave, then ask the decision maker to specify new marginal rates of substitution.

(7) Determine the step-size by solving the ε-constraint problem with the upper bounds $\varepsilon_j^h + t\Delta\varepsilon_j^h$, $j = 1, \ldots, k$, $j \neq \ell$, for different values of t. Denote the optimal value of the objective function by $z_\ell^h(t)$. A step-size t^h maximizing the proxy function is selected. If the new objective vector $(\varepsilon_1^h + t^h\Delta\varepsilon_1^h, \ldots, z_\ell^h(t^h), \ldots, \varepsilon_k^h + t^h\Delta\varepsilon_k^h)^T$ is preferred to $\mathbf{z}^h$, denote the corresponding decision vector by $\mathbf{x}^{h+1}$, set $h = h + 1$ and go to step (3). If the decision maker prefers $\mathbf{z}^h$ to the new solution, reduce t^h to be $\frac{1}{2}t^h, \frac{1}{4}t^h, \ldots$ until an improvement is achieved.

The maximum of the proxy function is determined by altering the step-size t, calculating the corresponding Pareto optimal solution and searching for three t values, t_1, t^h and t_2 so that $t_1 < t^h < t_2$ and $p(t_1) < p(t^h) > p(t_2)$, where p is the proxy function. When the condition above is satisfied, the local maximum of the proxy function $p(t)$ is in the neighbourhood of t^h.

Under assumptions 1–4 (in Subsection 5.3.1), the optimality condition for problem (5.3.1) at $\boldsymbol{\varepsilon}^h$ is that the gradient equals zero at that point. This means that $m_{\ell j}^h = \lambda_{\ell j}^h$ for $j = 1, \ldots k$, $j \neq \ell$. This is the background of the absolute value checking at step (4) (see also (2.8.1) in Part I).

5.3.3. Comments

The consistency of the marginal rates of substitution is checked because it is important for the successful convergence of the algorithm. The consistency at a single point is tested by the chain rule and by limiting the discrepancy (the formula is given in Sakawa (1982)) by a given tolerance level. The consistency at successive points is tested by checking the concavity and monotonicity of the proxy function (the proxy function must fulfill the same assumptions as the value function). A theorem giving conditions for different types of proxy functions is presented in Sakawa (1982).

To ensure the convergence of the algorithm it must, at each iteration, be checked that a sufficient improvement is obtained. If the decision maker prefers the new solution, the procedure may continue. Otherwise, a new step-size must be estimated.

It is remarked in Sakawa (1982) that the SPOT algorithm is nothing but a feasible direction method as for the convergence rate. The convergence can be demonstrated by the convergence of the modified feasible direction method. For this statement to be true, an ideal (i.e., consistent with correct answers) decision maker must be assumed.

SPOT can be classified among methods of a non ad hoc nature. If the value function is known, the marginal rates of substitution can be computed directly and the step-size is easy to calculate.

5.3.4. Applications and Extensions

The functioning of the SPOT algorithm is demonstrated in Sakawa (1982) by an academic example. It is shown that even though the marginal rates of substitution are only approximations, this does not necessarily worsen the results remarkably. A problem concerning industrial pollution in Osaka City in Japan is solved by SPOT in Sakawa and Seo (1980, 1982a, b). The problem is defined as a large-scale problem in Sakawa and Seo (1980) and a dual decomposition method is used to solve the ε-constraint problems.

A fuzzy SPOT is presented in Sakawa and Yano (1985). The decision maker is assumed to assess the marginal rates of substitution in a fuzzy form. In Sakawa and Mori (1983), a new method for nonconvex problems is proposed, where the weighted Tchebycheff problem is used to generate Pareto optimal solutions instead of the ε-constraint method, and trade-off rates are not used. A method related to the preceding one is presented in Sakawa and Mori (1984). The difference is a penalty scalarizing function used in generating Pareto optimal solutions (see Section 3.5). This method is also applicable to nonconvex problems.

5.3.5. Concluding Remarks

Ideas from several methods are combined in SPOT and several concepts are utilized. As far as the role of the decision maker is concerned, (s)he is only required to determine the marginal rates of substitution. Difficulties related to this determination were mentioned in Section 5.2 and they are still valid. However, the consistency of the marginal rates of substitution in SPOT is even more important than in the GDF method. This is a very demanding requirement.

A positive feature of SPOT when compared to the GDF method is that only Pareto optimal solutions are handled. Because the multiobjective optimization

problem was assumed to be convex, globally Pareto optimal solutions are obtained. The burden on the decision maker is decreased by employing a proxy function when selecting the step-size.

Many assumptions are set to guarantee the proper functioning of the algorithm. Some of these are quite difficult to check in practice (see concluding remarks concerning the GDF method in Subsection 5.2.5).

5.4. Tchebycheff Method

The Tchebycheff method, proposed in Steuer (1986, pp. 419–450) and Steuer and Choo (1983) and refined in Steuer (1989a), is an interactive weighting vector space reduction method. Originally, it was called the interactive weighted Tchebycheff procedure. A notable difference when compared to the methods described thus far is that a value function is not used in the Tchebycheff method. In addition, the role of the decision maker is different and somewhat simpler. Here, we introduce the Tchebycheff algorithm according to the refined version but modified for minimization problems.

5.4.1. Introduction

The Tchebycheff method has been designed to be user-friendly for the decision maker, and, thus, complicated information is not required. To start with, a utopian objective vector below the ideal objective vector is established. Then the distance from the utopian objective vector to the feasible objective region, measured by a weighted Tchebycheff metric, is minimized. Different solutions are obtained with different weighting vectors in the metric, as introduced in Section 3.4. The solution space is reduced by working with sequences of smaller and smaller subsets of the weighting vector space. Thus, the idea is to develop a sequence of progressively smaller subsets of the Pareto optimal set until a final solution is located. At each iteration, different alternative objective vectors are presented to the decision maker and (s)he is asked to select the most preferred of them. The feasible region is then reduced and alternatives from the reduced space are presented to the decision maker for selection.

Contrary to the previous interactive methods for multiobjective optimization, the Tchebycheff method does not presume many assumptions regarding the problem to be solved. It is assumed that

1. Less is preferred to more by the decision maker.
2. The objective functions are bounded (from below) over the feasible region S.

In what follows we assume that the global ideal objective vector and, thus, the global utopian objective vector are known, and we can leave the absolute value signs from the metrics. The metric to be used for measuring the distances

to a utopian objective vector is the weighted Tchebycheff metric (see Section 3.4). That is, the function to be minimized is

$$\max_{i=1,\ldots,k} \left[w_i(f_i(\mathbf{x}) - z_i^{**}) \right], \tag{5.4.1}$$

where $\mathbf{w} \in W = \{\mathbf{w} \in \mathbf{R}^k \mid 0 < w_i < 1, \ \sum_{i=1}^k w_i = 1\}$. We have a family of metrics since $\mathbf{w} \in W$ can vary widely. This nondifferentiable problem can be solved as a differentiable weighted Tchebycheff problem (3.4.3) (where the ideal objective vector is replaced by the utopian objective vector).

According to Theorem 3.4.5, we know that every Pareto optimal solution of any multiobjective optimization problem can be found by solving the weighted Tchebycheff problem with $\mathbf{z}^{**}$. The negative aspect with this problem is that some of the solutions may be weakly Pareto optimal. This weakness was handled in Subsection 3.4.5. Producing weakly Pareto optimal solutions is overcome in the Tchebycheff method by formulating the distance minimization problem as a *lexicographic weighted Tchebycheff problem*:

$$\begin{aligned} &\text{lex minimize} && \max_{i=1,\ldots,k} \left[w_i(f_i(\mathbf{x}) - z_i^{**}) \right], \sum_{i=1}^{k} (f_i(\mathbf{x}) - z_i^{**}) \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{5.4.2}$$

The functioning of problem (5.4.2) is described in Figure 5.4.1 by a problem with two objective functions. The bold line illustrates the Pareto optimal set. The weighted Tchebycheff problem has L-shaped contours (the thin continuous line) whose vertices lie along the line emanating from $\mathbf{z}^{**}$ in the direction $(1/w_1, 1/w_2, \ldots, 1/w_k)$. When minimizing the distance, a contour is determined which is closest to $\mathbf{z}^{**}$ and intersects Z. If this problem does not have a unique solution, that is, there are several feasible points on the optimal contour intersecting Z, then some of them may not be Pareto optimal. In practice, the uniqueness is usually difficult to check, and, to be on the safe side, the following step must be taken. In this case, the sum term is minimized subject to the obtained points to determine which of them is closest to $\mathbf{z}^{**}$ according to the L_1-metric (the dotted line). Thus a Pareto optimal solution (see Theorem 3.4.1) is obtained.

The following theorems formulate the connection between the lexicographic weighted Tchebycheff problem and Pareto optimal solutions.

Theorem 5.4.1. The solution of lexicographic weighted Tchebycheff problem (5.4.2) is Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a solution of problem (5.4.2). Let us assume that it is not Pareto optimal. In this case there exists some $\mathbf{x}^\circ \in S$ such that $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*)$ for at least one j. This and the positivity of the weights implies that $w_i(f_i(\mathbf{x}^\circ) - z_i^{**}) \leq w_i(f_i(\mathbf{x}^*) - z_i^{**})$ for every i and thus $\max_i[f_i(\mathbf{x}^\circ) - z_i^{**}] \leq \max_i[f_i(\mathbf{x}^*) - z_i^{**}]$.

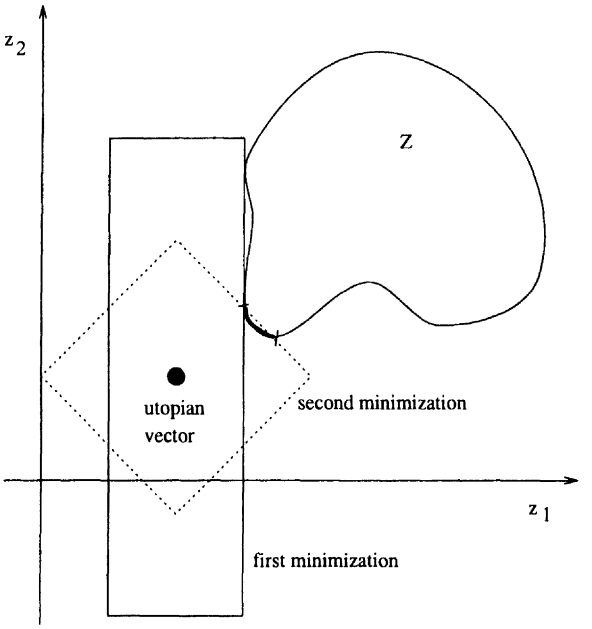

Figure 5.4.1. Lexicographic weighted Tchebycheff problem.

On the other hand, $f_i(\mathbf{x}^\circ) - z_i^{\star\star} \leq f_i(\mathbf{x}^*) - z_i^{\star\star}$ for all $i = 1, \ldots, k$ and at least one of the inequalities is strict. That is why we have $\sum_{i=1}^k (f_i(\mathbf{x}^\circ) - z_i^{\star\star}) < \sum_{i=1}^k (f_i(\mathbf{x}^*) - z_i^{\star\star})$. Here we have a contradiction with $\mathbf{x}^*$ being a solution of (5.4.2). Thus, $\mathbf{x}^*$ is Pareto optimal. □

Theorem 5.4.2. Let $\mathbf{x}^* \in S$ be Pareto optimal. Then there exists a weighting vector $\mathbf{0} < \mathbf{w} \in \mathbf{R}^k$ such that $\mathbf{x}^*$ is a unique solution of lexicographic weighted Tchebycheff problem (5.4.2).

Proof. Let $\mathbf{x}^* \in S$ be Pareto optimal. Let us assume that there exists no weighting vector $\mathbf{w} > \mathbf{0}$ such that $\mathbf{x}^*$ is a unique solution of problem (5.4.2).

We know that $f_i(\mathbf{x}) > z_i^{\star\star}$ for all $i = 1, \ldots, k$ and for all $\mathbf{x} \in S$. That is why we can choose for all $i = 1, \ldots, k$

$$w_i = \frac{1}{f_i(\mathbf{x}^*) - z_i^{\star\star}} \left(\sum_{i=1}^k \frac{1}{f_i(\mathbf{x}^*) - z_i^{\star\star}} \right)^{-1}.$$

If $\mathbf{x}^*$ is not a unique solution of (5.4.2), there exists another point $\mathbf{x}^\circ \in S$ that is a solution of this lexicographic weighted Tchebycheff problem. This implies that $\mathbf{x}^\circ$ must be a solution of the weighted Tchebycheff problem. This means that

$$\max_{i=1,\ldots,k}\left[\frac{1}{f_i(\mathbf{x}^*)-z_i^{**}}\left(\sum_{i=1}^{k}\frac{1}{f_i(\mathbf{x}^*)-z_i^{**}}\right)^{-1}(f_i(\mathbf{x}^\circ)-z_i^{**})\right]$$
$$\leq \max_{i=1,\ldots,k}\left[\frac{1}{f_i(\mathbf{x}^*)-z_i^{**}}\left(\sum_{i=1}^{k}\frac{1}{f_i(\mathbf{x}^*)-z_i^{**}}\right)^{-1}(f_i(\mathbf{x}^*)-z_i^{**})\right]$$
$$=\left(\sum_{i=1}^{k}\frac{1}{f_i(\mathbf{x}^*)-z_i^{**}}\right)^{-1}.$$

After simplifying the expression we have

$$f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$$

for every $i = 1, \ldots, k$. Because $\mathbf{x}^*$ is Pareto optimal, we must have $f_i(\mathbf{x}^\circ) = f_i(\mathbf{x}^*)$ for all i. In other words, the weighted Tchebycheff problem, and thus also the lexicographic weighted Tchebycheff problem, has a unique solution.

□

An alternative proof of Theorems 5.4.1 and 5.4.2 is given in Steuer (1986, p. 445) and Steuer and Choo (1983). Now we know that the lexicographic weighted Tchebycheff problem produces Pareto optimal solutions and any Pareto optimal solution can be found.

In the Tchebycheff method, different Pareto optimal solutions are obtained by altering the weighting vector. At each iteration h, the weighting vector space $W^h = \{\mathbf{w}^h \in \mathbf{R}^k \mid l_i^h < w_i^h < u_i^h,\ \sum_{i=1}^k w_i^h = 1\}$ is reduced to W^{h+1}, where $W^{h+1} \subset W^h$. With a sequence of progressively smaller subsets of the weighting vector space, a sequence of smaller subsets of the Pareto optimal set is sampled.

At the first iteration, a sample of the whole Pareto optimal set is generated by solving the lexicographic weighted Tchebycheff problem with well dispersed weighting vectors from $W = W^1$ (with $l_i^1 = 0$ and $u_i^1 = 1$). The reduction of W^h is done by tightening the upper and the lower bounds for the weighting vectors. Let $\mathbf{z}^h$ be the objective vector that the decision maker chooses from the sample at the iteration h and let $\mathbf{w}^h$ be the corresponding weighting vector in problem (5.4.2). Now a concentrated group of weighting vectors centred around $\mathbf{w}^h$ is formed. In this way, a sample of Pareto optimal solutions centred about $\mathbf{z}^h$ is obtained. It is advised to use normalized objective functions in the calculations.

The number of the alternative objective vectors to be presented to the decision maker is denoted by P. The number is usually specified by the decision maker. It may be fixed or different at each iteration. The algorithm becomes more reliable, if as many alternatives as possible can be evaluated effectively at each iteration. Human capabilities are yet limited, and some kind of a compromise is desirable.

When reducing the weighting vector space at each iteration, a reduction factor r is needed. The larger the reduction factor is, the faster the weighting vector space is reduced and the smaller are the decision maker's possibilities

for making errors and changing her or his mind concerning her or his desires during the process. The correct selection of r is thus important. It is suggested in Steuer (1986) and Steuer and Choo (1983) that

$$(1/P)^{1/k} \lessapprox r \lessapprox v^{1/(H-1)},$$

where v is the final interval length of the weighting vectors with $\frac{1}{2k} \lessapprox v \lessapprox \frac{3}{2k}$, H is the number of iterations to be carried out and $\lessapprox$ stands for 'approximately equal or less.'

5.4.2. Tchebycheff Algorithm

We can now present the main features of the Tchebycheff algorithm.

(1) Specify values for the set size $P(\gtrapprox k)$, a reduction factor $r < 1$ and an approximation for the number of iterations $H(\approx k)$. Set $l_i^1 = 0$ and $u_i^1 = 1$ for all $i = 1, \ldots, k$. Construct the utopian objective vector. Set $h = 1$.
(2) Form the weighting vector space $W^h = \{\mathbf{w}^h \in \mathbf{R}^k \mid l_i^h < w_i^h < u_i^h, \ \sum_{i=1}^k w_i^h = 1\}$.
(3) Generate $2P$ dispersed weighting vectors $\mathbf{w}^h \in W^h$.
(4) Solve lexicographic weighted Tchebycheff problem (5.4.2) for each of the $2P$ weighting vectors.
(5) Present the P most different of the resulting objective vectors to the decision maker and let her or him choose the most preferred among them, denoting it by $\mathbf{z}^h$.
(6) If $h = H$ go to step (8). Otherwise, modify, if necessary, the weighting vector corresponding to $\mathbf{z}^h$ such that if problem (5.4.2) was solved again, $\mathbf{z}^h$ would be a uniquely generated solution at the vertex of the intersecting new contour.
(7) Specify l_i^{h+1} and u_i^{h+1} for the reduced weighting vector space W^{h+1}, set $h = h + 1$ and go to step (2).
(8) The final solution is $\mathbf{x}^h$ corresponding to $\mathbf{z}^h$.

Dispersed weighting vectors are generated from W^h in step (3). In practice, this can be realized by generating randomly a large set (e.g., $50k$) of weighting vectors. Then the vectors are filtered (see Steuer (1986, pp. 311–326)) or clustered. The clustering is practical since subroutines for it are available in many subroutine libraries (such as IMSL). While we want to obtain $2P$ well dispersed weighting vectors, we form $2P$ clusters and choose one candidate from each of them either arbitrarily or near the centre.

Computationally, the following algorithm can be used to obtain random weighting vectors in W^h. We omit the index h for clarity.

For $i = 1, \ldots, k$ set

$$we_i = l_i + ra(u_i - l_i),$$

where ra is a random number between 0 and 1. Calculate the sums $we = \sum_{i=1}^{k} we_i$, $l = \sum_{i=1}^{k} l_i$ and $u = \sum_{i=1}^{k} u_i$. Finally, set for each $i = 1, \ldots, k$

$$w_i = \begin{cases} we_i + \dfrac{1 - we}{l - we}(l_i - we_i) & \text{if } we > 1, \\ we_i + \dfrac{1 - we}{u - we}(u_i - we_i) & \text{if } we < 1, \\ we_i & \text{otherwise.} \end{cases}$$

The lexicographic weighted Tchebycheff problem is solved $2P$ times for $2P$ weighting vectors (instead of P) to overcome the case that the same or a very similar solution is obtained with different weighting vectors. The $2P$ (or less) objective vectors are again filtered or clustered to obtain the P most different. In this way, it is guaranteed that P different objective vectors can be presented to the decision maker. For graphical illustration of the alternatives, see Chapter 3 in Part III.

The modification of the weighting vector in step (6) can be conducted by setting

$$w_i^h = \frac{1}{f_i(\mathbf{x}^h) - z_i^{\star\star}} \left(\sum_{i=1}^{k} \frac{1}{f_i(\mathbf{x}^h) - z_i^{\star\star}} \right)^{-1}$$

for all $i = 1, \ldots, k$ as in the proof of Theorem 5.4.2. The modification is relevant because the weighting vector space is reduced with respect to $\mathbf{w}^h$. It is useful to have an unbiased basis for the reduction.

Several possibilities for reducing the weighting vector space have been suggested. It is proposed in Steuer (1986) to set

$$(l_i^{h+1}, u_i^{h+1}) = \begin{cases} (0, r^h) & \text{if } w_i^h - \frac{r^h}{2} \leq 0, \\ (1 - r^h, 1) & \text{if } w_i^h + \frac{r^h}{2} \geq 1, \\ (w_i^h - \frac{r^h}{2}, w_i^h + \frac{r^h}{2}) & \text{otherwise,} \end{cases}$$

where r^h means raising r to the power h. In Steuer (1989a), an auxiliary scalar ω is determined so that the ratio of the volumes of W^{h+1} and W^h is r. Then ω is used in the reduction instead of the term $\frac{1}{2}r^h$.

The predetermined number of iterations is not necessarily conclusive. The decision maker can stop iterating when (s)he obtains a satisfactory solution or continue the solution process longer if necessary.

5.4.3. Comments

It is suggested in Steuer (1986, 1989a) that the sampling of the Pareto optimal set works in the most unbiased way if the ranges of the objective function values over the Pareto optimal set are approximately the same. This can be accomplished by re-scaling the objective functions in a way similar to that presented in Subsection 2.4.3 of Part I, when necessary. It is advisable

to use the scaling only in the calculations and present the alternatives to the decision maker in the original form. More suggestions for modifications of the algorithm are presented in Steuer (1989a).

The convergence rate of the Tchebycheff method is very difficult to establish. It is stressed in Steuer (1989a) that the Tchebycheff method is able to converge to any Pareto optimal solution. The reduction factor r is comprehended as a convergence factor because it determines how fast the reduction takes place. The weighting vector space is reduced until a solution is obtained that is satisfactory enough to be a final solution (see Steuer and Choo (1983)).

The Tchebycheff method can be characterized as a non ad hoc method. If the value function is known, it is easy to select from the set of P alternatives the one maximizing the value function.

We do not here go into details of the alternative version of the Tchebycheff method. We only mention that the possibility of getting weakly Pareto optimal solutions may be overcome by using augmented weighted Tchebycheff problem (3.4.5) (see Figure 3.4.2). This means that properly Pareto optimal solutions are handled instead of Pareto optimal ones (see Theorem 3.4.6). In this way, the lexicographic optimization is avoided, but the Tchebycheff algorithm is more complicated in other ways. For example, the determination of the correct value for the augmentation parameter ρ brings additional problems. It is proved in Steuer (1986, pp. 440–444) and Steuer and Choo (1983) that the augmented weighted Tchebycheff problem can be used to characterize Pareto optimal solutions if the feasible region is finite or all the constraints are linear. A numerical illustration of the algorithm is presented in Steuer (1986, pp. 468–472).

Implementing the Tchebycheff method in a spreadsheet (Excel) environment is suggested in Steuer (1997). The Tchebycheff method in its augmented form is applied in Wood et al. (1982) to water allocation problems of a river basin and in Silverman et al. (1988) to manpower supply forecasting. The augmented method form is also used in Agrell et al. (1998) when solving an MOLP problem of reservoir management. In Olson (1993), the Tchebycheff method is applied to a sausage blending problem and in Kaliszewski (1987) it is proposed that modified weighted Tchebycheff problem (3.4.6) is used to minimize the distances in the Tchebycheff method.

5.4.4. Concluding Remarks

A positive feature of the Tchebycheff method is that the role of the decision maker is quite easy to understand. (S)he does not need to realize new concepts or specify numerical answers as, for example, in the ISWT and the GDF methods. All (s)he has to do is to compare several alternative objective vectors and select the most preferred one. The ease of the comparison depends on the magnitude of P and on the number of objective functions. The personal capabilities of the decision makers also play an important role. It is also positive that all the alternatives are Pareto optimal.

The flexibility of the method is reduced by the fact that the discarded parts of the weighting vector space cannot be restored if the decision maker changes her or his mind. Thus, some consistency is required.

The weakness of the Tchebycheff method is that a great deal of calculation is needed at each iteration and many of the results are discarded. For large and complex problems, where the evaluation of the values of the objective functions may be laborious, the Tchebycheff method is not a realistic choice. On the other hand, it is possible to utilize parallel computing since all the lexicographic problems can be solved independently.

Although no absolute superiority can be attributed, it is worth mentioning that the Tchebycheff method performed best in the comparative evaluation of four methods (the ZW, the SWT, the Tchebycheff and the GUESS methods) in Buchanan and Daellenbach (1987) (see Subsection 1.2.3 of Part III). However, a difficulty was encountered in comprehending the information provided. The test example had only three objective functions and six alternatives were presented at each iteration. And the cognitive burden only becomes larger when the number of the objective functions is increased.

5.5. Step Method

The step method (STEM), presented in Benayoun et al. (1971), contains elements somewhat similar to the Tchebycheff method, but is based on a different idea. STEM is one of the first interactive methods developed for multiobjective optimization problems. It was originally designed for the maximization of MOLP problems but can be extended for nonlinear problems, as described, for example, in Eschenauer et al. (1990b) and Sawaragi et al. (1985, pp. 268–269). It can be considered to aspire at finding satisfactory solutions instead of optimizing an underlying value function. We describe the method for the minimization of nonlinear problems.

5.5.1. Introduction

It is assumed in STEM that at a certain Pareto optimal objective vector the decision maker can indicate both functions that have acceptable values and those whose values are too high. The latter can be said to be unacceptable. The decision maker is now assumed to allow the values of some acceptable objective functions to increase so that the unacceptable functions can have lower values. In other words, (s)he must give up a little in the value(s) of some objective function(s) f_i ($i \in I^>$) in order to improve the values of some other objective functions f_i ($i \in I^<$) such that $I^> \cup I^< = \{1, \ldots, k\}$.

STEM uses the weighted Tchebycheff problem (3.4.2) to generate new solutions. The ideal objective vector $\mathbf{z}^\star$ is used as a reference point in the calculations. According to Theorem 3.4.2 the solutions are weakly Pareto optimal.

It is assumed that

1. Less is preferred to more by the decision maker.
2. The objective functions are bounded over the feasible region S.

Information concerning the ranges of the Pareto optimal set is needed in determining the weighting vector for the metric. The idea is to make the scales of all the objective functions similar with the help of the weighting coefficients.

The nadir objective vector $\mathbf{z}^{\text{nad}}$ is approximated from the payoff table as explained in Subsection 2.4.2 of Part I. Thus, the maximal element of the column i is called z_i^{nad}. The weighting vector is calculated by the formula

$$w_i = \frac{e_i}{\sum_{j=1}^{k} e_j}, \quad i = 1, \ldots, k,$$

where for every $i = 1, \ldots, k$

$$e_i = \frac{1}{z_i^\star} \frac{z_i^{\text{nad}} - z_i^\star}{z_i^{\text{nad}}}$$

as suggested in Eschenauer et al. (1990b), or

$$e_i = \frac{z_i^{\text{nad}} - z_i^\star}{\max \left[|z_i^{\text{nad}}|, |z_i^\star| \right]}$$

as suggested in Vanderpooten and Vincke (1989). (The denominators are not allowed to be zero.) The weight is larger for those objective functions that are far from their ideal objective vector component.

5.5.2. STEM Algorithm

The basic phases of the STEM algorithm are the following:

(1) Calculate the ideal and the nadir objective vectors and the weighting coefficients. Set $h = 1$. Solve weighted Tchebycheff problem (3.4.2) with the calculated weights. Denote the solution by $\mathbf{x}^h \in S$ and the corresponding objective vector by $\mathbf{z}^h \in Z$.

(2) Ask the decision maker to classify the objective functions at $\mathbf{z}^h$ into satisfactory $I^>$ and unsatisfactory ones $I^<$. If the latter class is empty, go to step (4). Otherwise, ask the decision maker to specify relaxed upper bounds ε_i for the satisfactory objective functions.

(3) Solve problem (5.5.1), where the upper bounds are taken into account. Denote the solution by $\mathbf{x}^{h+1} \in S$ and the corresponding objective vector by $\mathbf{z}^{h+1} \in Z$. Set $h = h + 1$. Go to step (2).

(4) Stop. The final solution is $\mathbf{x}^h$.

In the first step the distance between the ideal objective vector and the feasible objective region is minimized by the weighted Tchebycheff metric (the weighting coefficients specified as above). The solution obtained is presented to

the decision maker. Then the decision maker is asked to specify those objective function(s) whose value(s) (s)he is willing to relax (i.e., weaken) to decrease the values of some other objective functions. (S)he must also determine the amount(s) of acceptable relaxation. Ways of helping the decision maker in this phase are presented in Benayoun et al. (1971).

The feasible region is restricted according to the information of the decision maker and the weights of the relaxed objective functions are set equal to zero, that is $w_i = 0$ for $i \in I^>$. Then a new distance minimization problem

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} \left[w_i | f_i(\mathbf{x}) - z_i^\star | \right] \\ &\text{subject to} && f_i(\mathbf{x}) \leq \varepsilon_i \quad \text{for all} \quad i \in I^>, \\ &&& f_i(\mathbf{x}) \leq f_i(\mathbf{x}^h) \quad \text{for all} \quad i \in I^<, \\ &&& \mathbf{x} \in S \end{aligned} \tag{5.5.1}$$

is solved. The first new constraint set allows the relaxed (acceptable) objective function values to increase up till the specified level and the second new constraint set makes sure that the unsatisfactory objective function values do not increase, that is, get worse. The procedure continues until the decision maker does not want to change any component of the current objective vector. If the decision maker is not satisfied with any of the components, then the procedure must also be stopped. In this case, STEM fails to find a satisfactory solution.

Different versions of the method vary in the formulation of the constraint set. In some versions, a new constraint set is generated at every iteration and in some other versions new constraints are included to accompany the old ones. In the latter model the decision maker must be somewhat consistent in her or his actions because it is not possible to withdraw the restrictions set on the feasible region.

5.5.3. Comments

STEM does not assume the existence of an underlying value function. Even if one were available, it would not help in answering the questions. Thus STEM can be characterized as an ad hoc method. Naturally, nothing can be said about the convergence of STEM with respect to a value function. However, the developers of the method mention that the algorithm produces a final solution fast if the new constraints constructed during the solution process become ineligible for further relaxations.

A linear numerical application example of STEM is given in Hwang and Masud (1979, pp. 174–182). The properties of the solution set of STEM are studied in Crama (1983). A so-called exterior branching algorithm is presented in Aubin and Näslund (1972). It is another kind of extension of STEM into nonlinear problems. There are several differences when compared with the original method. For example, the decision maker does not need to specify any amounts of change and an implicit value function is assumed to exist. Some

extensions and modifications of STEM are also mentioned in Chankong and Haimes (1983b, p. 329).

5.5.4. Concluding Remarks

Because we are moving around the (weakly) Pareto optimal set, a decrement in some objective function values can be achieved only by paying the price of an increment in some other objective function values. The idea of specifying objective functions whose values should be decreased or can be increased seems quite simple and appealing. However, it may be difficult to estimate appropriate amounts of increment that would allow the desired amount of improvement in those functions whose values should be decreased. In other words, the control of the solution is somewhat indirect. On the other hand, a positive feature is that the information handled is easy to understand. No complicated concepts are introduced to the decision maker.

According to the results presented in Section 3.4, the solutions of STEM are not necessarily Pareto optimal, but weakly Pareto optimal solutions may be obtained. It must also be kept in mind that the global ideal objective vector has to be known.

STEM was the first interactive method to be based on the classification idea. Numerous other methods adapting this idea in one way or the other have appeared since. In what follows, we present several methods where the decision maker can specify both the amounts of relaxation and desirable aspiration levels. In this way the decision maker can control the solution process in a more direct way than in STEM.

5.6. Reference Point Method

As its name suggests, the reference point method, presented in Wierzbicki (1980a, b, 1981, 1982), is based on a reference point of aspiration levels. The reference point is a feasible or infeasible point in the objective space which is reasonable or desirable to the decision maker. The reference point is used to derive achievement scalarizing functions having minimal solutions at weakly, ε-properly or Pareto optimal points as introduced in Section 3.5. In this method, generating Pareto optimal solutions is based on reference points, not on value functions or weighting vectors. No specific assumptions are set on the problem to be solved. The reference point idea has been utilized in several methods in different ways. Wierzbicki's reference point method was among the first of them.

5.6.1. Introduction

The basic idea behind the reference point method is to reconsider how decision makers make decisions. It is doubted in Wierzbicki (1980a, b) that individuals make everyday decisions by maximizing a certain value function. Instead, Wierzbicki claims that decision makers want to attain certain aspiration levels (e.g., when making purchases according to a shopping list). He suggests that, while thousands of consumers may behave on the average as if they were maximizing a value function, no individual behaves in that way. The basic idea is satisficing (introduced in Section 2.6 of Part I) rather than optimizing. In addition, reference points are intuitive and easy for the decision maker to specify and their consistency is not an essential requirement.

Classifying the objective functions into acceptable and unacceptable ones (at a current objective vector) was mentioned in connection with STEM. Specifying a reference point can be considered a way of classifying the objective functions. If the aspiration level is lower than the current objective value, that objective function is currently unacceptable, and if the aspiration level is equal to or higher than the current objective value, that function is acceptable. The difference here is that the reference point can be infeasible in every component. In other words, where the set of acceptable objective functions is empty, the reference point-based approach can still be utilized. Naturally, this does not mean that all the objective values could be decreased but a different solution can be generated.

Further information concerning the matters addressed in this section can be found in Wierzbicki (1977, 1980b, 1981, 1982, 1986a, b). By a reference point method we here mean that of Wierzbicki's. The reference point method relies heavily on the properties of achievement functions, which were dealt with in Section 3.5. Of particular interest are Corollary 3.5.6 and Theorem 3.5.7. As far as the preference structure of the decision maker is concerned, it is assumed that

1. Less is preferred to more by the decision maker.

5.6.2. Reference Point Algorithm

The interactive multiobjective optimization technique of Wierzbicki is very simple and practical. Before the solution process starts, some information is given to the decision maker about the problem. If possible, the ideal objective vector and the (approximated) nadir objective vector are presented to illustrate the ranges of the Pareto optimal set. Another possibility is to minimize and maximize the objective functions individually in the feasible region (if it is bounded). Both decision variable and objective values are presented. An appropriate form for the achievement function must also be selected.

The basic steps of the reference point method are the following:

(1) Present information about the problem to the decision maker. Set $h = 1$.

(2) Ask the decision maker to specify a reference point $\bar{\mathbf{z}}^h \in \mathbf{R}^k$ (an aspiration level for every objective function).

(3) Minimize the achievement function and obtain a (weakly, ε-properly or) Pareto optimal solution $\mathbf{x}^h$ and the corresponding $\mathbf{z}^h$. Present $\mathbf{z}^h$ to the decision maker.

(4) Calculate a number of k other (weakly, ε-properly or) Pareto optimal solutions by minimizing the achievement function with perturbed reference points

$$\bar{\mathbf{z}}(i) = \bar{\mathbf{z}}^h + d^h \mathbf{e}^i,$$

where $d^h = \|\bar{\mathbf{z}}^h - \mathbf{z}^h\|$ and $\mathbf{e}^i$ is the ith unit vector for $i = 1, \ldots, k$.

(5) Present the alternatives to the decision maker. If (s)he finds one of the $k+1$ solutions satisfactory, the corresponding $\mathbf{x}^h$ is the final solution. Otherwise, ask the decision maker to specify a new reference point $\bar{\mathbf{z}}^{h+1}$. Set $h = h + 1$ and go to step (3).

The reason for writing the words weakly or ε-properly in parentheses in the algorithm is that it depends on the achievement function selected whether the solutions are weakly, ε-properly or Pareto optimal.

The advantage of perturbing the reference point in step (4) is that the decision maker gets a better conception of the possible solutions. If the reference point is far from the Pareto optimal set, the decision maker gets a wider description of the Pareto optimal set and if the reference point is near the Pareto optimal set, then a finer description of the Pareto optimal set is given. The effects of the perturbation and close and distant reference points are illustrated in Figure 5.6.1.

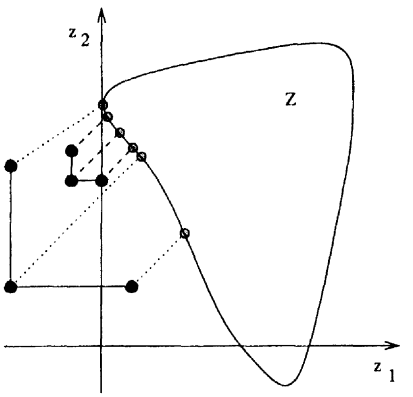

Figure 5.6.1. Altering the reference points.

5.6.3. Comments

As to the infinite convergence of the algorithm, the following result is stated in Wierzbicki (1980b).

Theorem 5.6.1. If the solutions of the achievement function in the algorithm are unique and if the minimal value of $\|\mathbf{z} - \bar{\mathbf{z}}\|$ subject to Pareto optimal objective vectors is equal to the minimal value of the achievement function $s_{\bar{\mathbf{z}}}$ subject to Z for $\bar{\mathbf{z}} \notin Z + \mathbf{R}^k_+$, then for any metric in $\mathbf{R}^k$, the solution procedure is convergent. In other words, $\lim_{h\to\infty} \|\mathbf{z}^h - \mathbf{z}^{h+1}\| = 0$.

Proof. See references in Wierzbicki (1980b).

A modification of the algorithm guaranteeing the convergence is presented in Wierzbicki (1980b).

The reference point method can be characterized as an ad hoc method or a method having both non ad hoc and ad hoc features. Alternatives are easy to compare if the value function is known. On the other hand, a reference point cannot be directly defined with the help of the value function. However, it is possible to test whether a new reference point has a higher value function value than the earlier solutions.

A different way of generating new reference points is suggested in Wierzbicki (1997b). It is a way of realizing the idea of a *reference ball* where a set of additional reference points in a ball of a fixed radius centered on the current solution is produced.

An appendix to the reference point method is suggested in Wierzbicki (1997b). After the decision maker has found a final solution (s)he can check whether more satisfactory solutions exist by a so-called outranking trials method. In the spirit of outranking methods of (discrete) multiattribute decision analysis, the decision maker is asked to specify preference, indifference and veto thresholds (see Subsection 5.9.1) for each objective function. Different states of outranking relations are established and a sequential questioning procedure is gone through with the decision maker to check whether there exist objective vectors whose components outrank the current final solution. This procedure may involve a lot of questions. The convergence of many other interactive methods can be investigated by the outranking trials method, as well.

5.6.4. Implementation

A software family called DIDAS (Dynamic Interactive Decision Analysis and Support) has been developed on the basis of the reference point ideas of Wierzbicki. The nonlinear version of DIDAS has been created and developed in several phases. For example, the International Institute for Applied Systems

Analysis (IIASA) in Austria and the Warsaw Technical University have been involved. The latest version is called IAC-DIDASN++. There is a lot of literature describing the various phases in the development work (see Granat et al. (1994a, b), Grauer (1983a, b), Grauer et al. (1984), Kreglewski (1989), Kreglewski et al. (1987, 1991), Lewandowski and Grauer (1982), Lewandowski et al. (1987) and Rogowski et al. (1987)).

DIDAS is a dynamic decision support system which aims at helping to achieve better decisions. The ideology has been extended from the reference point method with reservation levels. *Reservation levels* $\check{z}_i$ are objective function values the user wants to avoid. For the objective functions to be minimized they must be above the aspiration levels forming the reference point $\bar{\mathbf{z}}$. In DIDAS, the user is asked to specify both aspiration and reservation levels for each objective function. The achievement function has to be reformulated to take the reservation levels into account. Several achievement functions have been suggested in different versions of the system.

The user can easily obtain different Pareto optimal solutions by changing the aspiration levels and the reservation levels. The objective functions are scaled and the user is assumed to specify aspiration levels between the ideal objective vector and the nadir objective vector. In this setting, the user can implicitly attach more importance to attaining a particular aspiration level by placing it near the ideal objective value. In that case, the corresponding objective function is weighted stronger in the achievement function.

We give an example of achievement functions, including both aspiration and reservation levels. If all the objective functions are to be minimized, an order-approximating achievement function to be maximized can be of the form

$$\min_{i=1,\ldots,k}\left[\min\left[\frac{\check{z}_i - f_i(\mathbf{x})}{\check{z}_i - \bar{z}_i}, 1 + \frac{\nu(\bar{z}_i - f_i(\mathbf{x}))}{\bar{z}_i - z_i^\star}\right]\right] + \rho\left(\sum_{i=1}^{k}\min\left[\frac{\check{z}_i - f_i(\mathbf{x})}{\check{z}_i - \bar{z}_i}, 1 + \frac{\nu(\bar{z}_i - f_i(\mathbf{x}))}{\bar{z}_i - z_i^\star}\right]\right),$$

where $z_i^\star$ are components of the ideal objective vector, $\rho > 0$ is an augmentation term and

$$\nu = \min_{i=1,\ldots,k}\frac{\bar{z}_i - z_i^\star}{\check{z}_i - \bar{z}_i}.$$

Achievement functions may be computationally complicated. Further, they are typically nondifferentiable. However, their simplified and differentiable counterparts are generally used (see, e.g., Granat et al. (1994a) and Kreglewski et al. (1991)).

In further developing DIDAS, attention has been paid to computational efficiency in both achievement functions and their nonlinear single objective solvers. Gradient-based solvers are efficient and robust enough to be employed in interactive decision support systems. However, it is not advisable to ask for gradient information from the user of the system. Firstly, the formulation of the

derivatives is a time-consuming and laborious task, and secondly, errors and mistakes are likely to occur. (Mistakes have been found to be a main reason for the failure of nonlinear optimization methods in convergence.) The difficulties can be overcome with symbolic differentiation. This is briefly handled in Kreglewski (1989). One more alternative is to use automatic differentiation (see, e.g., Rall (1981)). However, no results of doing this have been reported.

DIDAS is general and can thus handle objective functions needing to be minimized, maximized or stabilized (i.e., the objective function should have a value as close to the given level as possible). Different objective function types imply changes in the achievement function used (see, for example, Granat et al. (1994a)).

5.6.5. Applications and Extensions

The reference point method is applied to econometric models in Olbrisch (1986). Some experiences in applying DIDAS to macroeconomics planning are reported in Grauer et al. (1984). DIDAS is used in empirical tests in Bischoff (1985) to experiment with different scalarizing functions. A problem of determining the optimal temperature in a greenhouse is solved by DIDAS in Udink ten Cate (1985). In Stam et al. (1992), DIDAS is used in analysing the acid rain problem in Europe. A trajectory-oriented extension of DIDAS is described and applied in Lewandowski et al. (1985a, b). Three applications of IAC-DIDASN++ for engineering design are reported in Wierzbicki and Granat (1997). They handle the design of a spur gear transmission unit, ship navigation support and automatic control.

The reference point idea is modified in Mocci and Primicerio (1997) to better handle nonconvex problems and avoid local optima. At each iteration, the achievement function is minimized in a reduced feasible region determined by the decision maker. This modified method is applied to a problem of ring network design.

The reference point method is generalized for several decision makers or several reference points in Song and Cheng (1988) and Vetschera (1991a). The reference point method is also essential in a group decision support system, described in Vetschera (1991b), where each group member uses the reference point method.

A so-called combined procedure combining the Tchebycheff method and the reference point method is introduced in Steuer et al. (1993). There the decision maker is asked to specify both reference points and the most satisfactory solution among the alternatives produced by the means of the two methods. So-called reference sets, extensions of reference points, are the basis of an interactive procedure described in Skulimowski (1996).

An extension of the reference point method, called the preemptive reference point method, is introduced in Ogryczak (1997a). The approach formulates reference point problems in the form of goal programming. Instead of considering

all the deviations to be equally important (which is the case in the reference point method) predefined priorities between the goals are also handled. The reference point method is modified for problems with homogeneous and anonymous objective functions in Ogryczak (1997b). Here, anonymity stands for symmetry with respect to permutations of the objective functions.

5.6.6. Concluding Remarks

Wierzbicki's reference point method is quite easy for the decision maker to understand. The decision maker only has to specify appropriate aspiration levels and compare objective vectors. What has been said about the comparison of alternatives in connection with the previous methods is also valid here. The solutions are weakly, ε-properly or Pareto optimal depending on the achievement function employed.

The freedom of the decision maker has both positive and negative aspects. The decision maker can direct the solution process and is free to change her or his mind during the process. However, the convergence is not necessarily fast if the decision maker is not purposeful. There is no clear strategy to produce the final solution since the method does not help the decision maker to find improved solutions.

Wierzbicki's method can be regarded as a generalization of goal programming. Aspiration levels are central in both methods, but unlike goal programming Wierzbicki's method is able to handle both feasible and infeasible aspiration levels.

Methods based on reference points are widely regarded efficient for the solution of practical problems. They are easy to understand and to implement. Further, they do not necessitate consistency from the decision maker. One can say that controlling a method with reference points is a more direct and a more explicit way than, for example, with weighting coefficients.

5.7. GUESS Method

The GUESS method is a simple interactive method related to the reference point method. The method is also sometimes called a naïve method and it is presented in Buchanan (1997). It will be referred to in Subsection 1.2.3 of Part III when describing method comparisons available in the literature. For this reason we present the method here briefly (modified for minimization problems).

5.7.1. Introduction

The GUESS method does not involve any special assumptions. The only requirement is that the ideal objective vector $\mathbf{z}^\star$ and the nadir objective vector $\mathbf{z}^{\text{nad}}$ are available. Thus, it is assumed that

1. Less is preferred to more by the decision maker.
2. The objective functions are bounded over the feasible region S.

The method proceeds as follows. The decision maker specifies a reference point (or a guess) $\bar{\mathbf{z}}^h$ and a solution with equal proportional achievements is generated. Then the decision maker specifies a new reference point and the iteration continues until the decision maker is satisfied with the solution produced. The search procedure is not assisted in any other way.

The scales of the objective functions are normalized with denominators $z_i^{\text{nad}} - z_i^\star$ for every $i = 1, \ldots, k$. The general idea is to maximize the minimum weighted deviation from the nadir objective vector. Thus, the idea is opposite to, for example, the weighted Tchebycheff problem where the maximum weighted deviation from the ideal objective vector is minimized.

We can put the same reasoning in other words. We can say that the objective functions are rescaled so that they all have the range $[0, 1]$. This means that each objective function $f_i(\mathbf{x})$ is replaced by a normalized function

$$\frac{z_i^{\text{nad}} - f_i(\mathbf{x})}{z_i^{\text{nad}} - z_i^\star} \quad \text{for all} \quad i = 1, \ldots, k.$$

Let us once again emphasize that the global ideal objective vector and the nadir objective vector are assumed to be known.

The *weighted max-min problem* to be solved is

$$\begin{aligned} &\text{maximize} && \min_{i=1,\ldots,k} \left[\frac{1}{w_i} \frac{z_i^{\text{nad}} - f_i(\mathbf{x})}{z_i^{\text{nad}} - z_i^\star} \right] \\ &\text{subject to} && \mathbf{x} \in S, \end{aligned} \tag{5.7.1}$$

where the weighting coefficients w_i, $i = 1, \ldots, k$, are positive and the denominators must not equal zero.

We have the following result.

Theorem 5.7.1. The solution of weighted max-min problem (5.7.1) is weakly Pareto optimal.

Proof. Let $\mathbf{x}^* \in S$ be a solution of the weighted max-min problem. Let us suppose that $\mathbf{x}^*$ is not weakly Pareto optimal. In this case, there exists a point $\mathbf{x} \in S$ such that $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$ for every $i = 1, \ldots, k$. This means that, $z_i^{\text{nad}} - f_i(\mathbf{x}) > z_i^{\text{nad}} - f_i(\mathbf{x}^*)$ for all i. While we have $w_i > 0$ and $z_i^{\text{nad}} - z_i^\star > 0$, we can write

$$\frac{1}{w_i}\frac{z_i^{\text{nad}} - f_i(\mathbf{x})}{z_i^{\text{nad}} - z_i^\star} > \frac{1}{w_i}\frac{z_i^{\text{nad}} - f_i(\mathbf{x}^*)}{z_i^{\text{nad}} - z_i^\star} \quad \text{for every } i = 1, \ldots, k.$$

This implies that

$$\min_{i=1,\ldots,k}\left[\frac{1}{w_i}\frac{z_i^{\text{nad}} - f_i(\mathbf{x})}{z_i^{\text{nad}} - z_i^\star}\right] > \min_{i=1,\ldots,k}\left[\frac{1}{w_i}\frac{z_i^{\text{nad}} - f_i(\mathbf{x}^*)}{z_i^{\text{nad}} - z_i^\star}\right].$$

Thus, $\mathbf{x}^*$ cannot be a solution of the weighted max-min problem. This contradiction completes the proof and $\mathbf{x}^*$ is weakly Pareto optimal. □

The weighting coefficients are not any positive numbers whatsoever, but normalized aspiration levels. In other words, we have

$$w_i^h = \frac{z_i^{\text{nad}} - \bar{z}_i^h}{z_i^{\text{nad}} - z_i^\star} \quad \text{for all } i = 1, \ldots, k.$$

With the specified weighting coefficients we can write the problem to be solved in the form

$$\begin{aligned} &\text{maximize} \quad \min_{i=1,\ldots,k}\left[\frac{z_i^{\text{nad}} - f_i(\mathbf{x})}{z_i^{\text{nad}} - \bar{z}_i^h}\right] \\ &\text{subject to} \quad \mathbf{x} \in S. \end{aligned} \tag{5.7.2}$$

Notice that the aspiration levels specified by the decision maker have to be strictly lower than the nadir objective vector, that is, $\bar{\mathbf{z}}^h < \mathbf{z}^{\text{nad}}$. If all the objective functions are differentiable, problem (5.7.2) can be written in a differentiable form with the help of an additional variable, whereas the nondifferentiable formulation can be solved with appropriate single objective optimizers.

We can prove that any Pareto optimal solution can be found with problem (5.7.2).

Theorem 5.7.2. If $\mathbf{x}^* \in S$ is Pareto optimal, then it is a solution of problem (5.7.2) with $\bar{\mathbf{z}} = \mathbf{f}(\mathbf{x}^*)$.

Proof. Let $\mathbf{x}^* \in S$ be Pareto optimal and let us suppose that it is not a solution of (5.7.2) with $\bar{\mathbf{z}} = \mathbf{f}(\mathbf{x}^*)$. In this case there exist another $\mathbf{x}^\circ \in S$ such that

$$\min_{i=1,\ldots,k}\left[\frac{z_i^{\text{nad}} - f_i(\mathbf{x}^\circ)}{z_i^{\text{nad}} - f_i(\mathbf{x}^*)}\right] > \min_{i=1,\ldots,k}\left[\frac{z_i^{\text{nad}} - f_i(\mathbf{x}^*)}{z_i^{\text{nad}} - f_i(\mathbf{x}^*)}\right] = 1.$$

This means that $f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*)$ for every $i = 1, \ldots, k$ which is a contradiction with the Pareto optimality of $\mathbf{x}^*$. In other words, $\mathbf{x}^*$ must be a solution of (5.7.2). □

According to Theorems 5.7.1 and 5.7.2 we know that all the solutions generated are weakly Pareto optimal and any Pareto optimal solution can be found.

5.7.2. GUESS Algorithm

The GUESS method has five basic steps.

(1) Calculate the ideal objective vector and the nadir objective vector and present them to the decision maker. Set $h = 1$.
(2) Let the decision maker specify upper or lower bounds to the objective functions if (s)he so desires. Update problem (5.7.2), if necessary.
(3) Ask the decision maker to specify a reference point $\bar{\mathbf{z}}^h \in \mathbf{R}^k$ between the ideal and the nadir objective vectors.
(4) Solve problem (5.7.2) and obtain a weakly Pareto optimal solution $\mathbf{x}^h$. Present the corresponding objective vector $\mathbf{z}^h$ to the decision maker.
(5) If the decision maker is satisfied with $\mathbf{z}^h$, set $\mathbf{x}^h$ as a final solution and stop. Otherwise, set $h = h + 1$ and go to step (2).

In step (2) the specification of upper or lower bounds means adding constraints to problem (5.7.2). Nevertheless, the components of the ideal or the nadir objective vector needed in the calculation elsewhere are not changed or affected.

The only stopping rule is the satisfaction of the decision maker. No guidance is given to the decision maker in setting new aspiration levels. This is typical of many reference point-based methods.

5.7.3. Comments

The GUESS method is based on trial and error. The decision maker can examine what kind of an effect her or his input has on the solution obtained and then modify the input, if necessary. The system does not provide any additional or supporting information about the problem to be solved.

As long as no additional constraints are included in the problem, the components of the solution obtained are in equal proportion with the components of the reference point specified. In other words, when the solution obtained and the corresponding reference point are normalized, the quotients of their component are the same for each component. The reason for this behaviour is that the reference point is contained in the weighting vector.

The GUESS method is an ad hoc method. The existence of a value function would not help in determining new reference points or upper or lower bounds for the objective functions.

An interesting practical observation is mentioned in Buchanan (1997). Namely, decision makers are easily satisfied if there is a small difference between the reference point and the solution obtained. Somehow they feel a need to be satisfied when they have almost achieved what they wanted. In this case they may stop iterating 'too early.' The decision maker is naturally allowed to stop the solution process if the solution really is satisfactory. But, the coincidence of setting the reference point near an attainable solution may unnecessarily increase the decision maker's satisfaction.

5.7.4. Concluding Remarks

The GUESS method is simple to use and does not set any specific assumptions on the behaviour or the preference structure of the decision maker. The decision maker can change her or his mind since no consistency is required. The only information required from the decision maker is a reference point and possible upper and lower bounds.

The method has been compared to several other interactive methods in different comparative evaluations (to be described in Subsection 1.2.3 of Part III). It has been received relatively well in the experiments reported. The reasons may be its simplicity and flexibility.

The optional upper or lower bounds specified by the decision maker are not checked in any way in the method. Inappropriate lower bounds may lead into solutions that are not weakly Pareto optimal. In other words, additional constraints may invalidate the result of Theorem 5.7.1. This can be avoided, for example, by allowing only upper bounds.

The weakness of the GUESS method is its heavy reliance on the availability of the nadir objective vector. As mentioned in Subsection 2.4.2 of Part I, the nadir objective vector is not easy to determine and it is usually only an approximation.

5.8. Satisficing Trade-Off Method

The satisficing trade-off method (STOM), presented in Nakayama (1989, 1995), Nakayama and Furukawa (1985), Nakayama and Sawaragi (1984) and Nakayama et al. (1986), is based on ideas similar to the reference point method of Wierzbicki and the GUESS method. The differentiating factor is the trade-off information utilized. The method is here presented according to Nakayama (1995).

5.8.1. Introduction

STOM originates from classification and aspiration levels. It is based on satisficing decision making, as can be deduced from its name.

The functioning of STOM is the following. After a weakly or a properly Pareto optimal solution has been obtained by optimizing a scalarizing function, it is presented to the decision maker. On the basis of this information (s)he is asked to classify the objective functions into three classes. The classes are the unacceptable objective functions whose values (s)he wants to improve ($I^<$), the acceptable objective functions whose values (s)he agrees to relax (impair) ($I^>$) and the acceptable objective functions whose values (s)he accepts as they are ($I^=$) (such that $I^< \cup I^> \cup I^= = \{1, \ldots, k\}$). Trade-off rate information is utilized so that the decision maker only has to specify aspiration levels for the

functions in $I^<$. Next, a modified scalarizing function is minimized and the decision maker is asked to classify the objective functions at the new solution.

The theoretical derivation of the method is based on the three requirements for scalarizing functions presented at the beginning of this part in Chapter 1. As was then stated, there does not exist a scalarizing function that can satisfy all three requirements. Nonetheless, a rather promising approach is to use the weighted Tchebycheff problem that can find any Pareto optimal solution. Hence it and its augmented variant are used in STOM.

Different forms of scalarizing functions have been suggested for use in STOM. In the original formulation the weighting coefficients are set as

$$w_i^h = \frac{1}{\bar{z}_i^h - z_i^{\star\star}} \quad \text{for every} \quad i = 1, \ldots, k, \tag{5.8.1}$$

where $\bar{\mathbf{z}}^h$ is a reference point and $\mathbf{z}^{\star\star}$ is a utopian objective vector so that $\bar{\mathbf{z}}^h > \mathbf{z}^{\star\star}$, and the scalarizing function to be minimized is once again (5.4.1).

If weakly Pareto optimal solutions are to be avoided, the scalarizing function used is

$$\max_{i=1,\ldots,k} [\, w_i^h (f_i(\mathbf{x}) - z_i^{\star\star}) \,] + \rho \sum_{i=1}^{k} w_i^h f_i(\mathbf{x}), \tag{5.8.2}$$

where ρ is some sufficiently small positive scalar, for example, of the order 10^{-6}. Both these scalarizing functions presume that the ideal objective vector and, thus, the utopian objective vector are known globally. However, if some objective function f_j is not bounded from below in S, then some small scalar value can be selected as $z_j^{\star\star}$.

If the problem is bounded, then the solutions obtained by function (5.4.1) are guaranteed to be weakly Pareto optimal (see Theorem 3.4.2) and every Pareto optimal solution can be found (see Theorem 3.4.5). Further, it is proved in Nakayama (1985a) and Sawaragi et al. (1985, pp. 271–272) that the solution obtained is satisficing (i.e., $f_i(\mathbf{x}^*) \leq \bar{z}_i^h$ for all $i = 1, \ldots, k$) if the reference point is feasible and weighting coefficients (5.8.1) are employed. For function (5.8.2) all the solutions are properly Pareto optimal and any properly Pareto optimal solution can be found. Even though the formulation slightly differs from (3.4.5), the results of Theorem 3.4.6 are still valid. Unfortunately, function (5.8.2) does not satisfy the third requirement concerning satisficing decision making (see Nakayama (1985a)).

Other forms of weighting coefficients can also be used. The selection affects the results obtained. This is demonstrated in Nakayama (1995). The reference point method-type achievement functions can be used as well. This means that the utopian objective vector is replaced by the reference point.

Both the scalarizing functions mentioned are nondifferentiable but they can be written in a differentiable form assuming the differentiability of the functions involved. This is carried out by introducing a scalar variable α as in (3.4.3).

In what follows, we refer to the differentiable form where all the objective functions have been transformed into constraints.

As mentioned in Subsection 3.4.4, trade-off rate information can be obtained with the help of differentiable formulation (3.4.3). Both weighting coefficients and Karush-Kuhn-Tucker multipliers are then utilized. That is why it must be assumed that

1. Less is preferred to more by the decision maker.
2. The objective and the constraint functions are twice continuously differentiable.

The availability of trade-off rates also necessitates the fulfillment of other assumptions mentioned in Subsection 3.4.4. They are parallel to those in Theorem 3.2.13; see also Yano and Sakawa (1987). This fact has not earlier been sufficiently emphasized when introducing the method.

5.8.2. STOM Algorithm

Let us now write down the steps of the algorithm.

(1) Calculate the utopian objective vector $\mathbf{z}^{\star\star}$. Set $h = 1$.
(2) Ask the decision maker to specify a reference point $\bar{\mathbf{z}}^h \in \mathbf{R}^k$ such that $\bar{z}_i^h > z_i^{\star\star}$ for every $i = 1, \ldots, k$.
(3) Minimize the scalarizing function used. Denote the solution by $\mathbf{x}^h$. Let the corresponding objective vector be $\mathbf{z}^h$. Present it to the decision maker.
(4) Ask the decision maker to classify the objective functions into the classes $I^<$, $I^>$ and $I^=$. If $I^< = \emptyset$, go to step (6). Otherwise, ask the decision maker to specify new aspiration levels $\bar{z}_i^{h+1}$ for the functions in $I^<$. Set $\bar{z}_i^{h+1} = f_i(\mathbf{x}^h)$ for $i \in I^=$.
(5) Let $\boldsymbol{\lambda}^h \in \mathbf{R}^k$ be the Karush-Kuhn-Tucker multipliers connected to $\mathbf{x}^h$. Use automatic trade-off to obtain new levels (upper bounds) $\bar{z}_i^{h+1}$ for the functions in $I^>$. Set $h = h + 1$ and go to step (3).
(6) Stop with the final solution $\mathbf{x}^h$.

It is naturally possible that the decision maker will also specify new levels for those objective functions whose values (s)he agrees to relax (i.e., increase). But particularly when the number of objective functions is great, the decision maker may appreciate the automatic trade-off feature when (s)he does not have to specify new aspiration levels for all the functions at each iteration. Naturally, the decision maker can modify the calculated aspiration levels if they are not agreeable.

As long as trade-off rates are obtainable from the Karush-Kuhn-Tucker multipliers and the weighting coefficients, the burden set on the decision maker can be decreased by employing *automatic trade-off* in specifying the aspiration levels (upper bounds) for the functions to be relaxed. They are derived

from sensitivity analysis on the basis of staying in the Pareto optimal set (see Nakayama (1991b, 1992a, 1995)). We set for each $i \in I^>$

$$\bar{z}_i^{h+1} = f_i(\mathbf{x}^h) + \frac{1}{N(\lambda_i^h + \rho)w_i^h} \sum_{j \in I^<} (\lambda_j^h + \rho) w_j^h (f_i(\mathbf{x}^h) - \bar{z}_i^{h+1}),$$

where N is the number of the objective functions in the class $I^>$. If no augmentation term is used in the scalarizing function, we set $\rho = 0$ in the formula above. Automatic trade-off increases all the objective functions in $I^>$ in the equal proportion to $(\lambda_i^h + \rho)w_i^h$. If the amounts of change are large or the problem is nonlinear, the aspiration levels produced by automatic trade-off may not be large enough to allow the desired improvements to the other objective functions (see Nakayama (1992b)).

5.8.3. Comments

If the problem is linear or quadratic, we can go even further than the automatic trade-off. In this case parametric optimization is used in generating so-called *exact trade-off*. This means that we can calculate exactly how much the objective function values must be relaxed in order to stay in the Pareto optimal set. Thus, we get a new Pareto optimal solution without having to re-optimize the scalarizing function (see Nakayama (1991b, 1992a, b)).

Trade-off information can also be used to check the feasibility of the reference point specified by the decision maker. If it is not feasible, the number of minimizations of the scalarizing function can be reduced by directly specifying higher aspiration levels (remember that satisficing solutions are obtained when the reference point is feasible in scalarizing function (5.4.1)). See details in Nakayama (1985a, 1989), Nakayama and Furukawa (1985) and Nakayama and Sawaragi (1984).

Trade-off information is valuable even if some Karush-Kuhn-Tucker multipliers are equal to zero. For example, if all the Karush-Kuhn-Tucker multipliers of the functions to be relaxed equal zero, we know that it is not possible to improve the desired objective function values with this classification. In other words, the functions to be relaxed cannot compensate for the improvement desired. The reason is that the objective functions to be relaxed are positively affected by other objective function(s) to be improved (see Nakayama (1995)).

Note that STOM can be used even in the absence of trade-off rate information. This may be the case if all the differentiability and the regularity assumptions are not satisfied. If trade-off rates are not used, no special assumptions need to be set on the problem to be solved. In this form STOM is almost the same as the GUESS method – only the achievement function used is different.

Because no specific assumptions are set on the underlying value function, convergence results based on it are not available. Even if a value function existed, it could not be directly used to determine the functions to be decreased

and increased or the amounts of change. Thus the method is characterized as an ad hoc method. On the other hand, one must remember that the aim of the method is particularly in satisficing rather than optimizing some value function.

It is pointed out in Nakayama (1995) and Nakayama et al. (1995) that the roles of the objective and the constraint functions can easily be interchanged when the nondifferentiable scalarizing function is solved in differentiable form (3.4.3). This is carried out by adding constant multipliers β_i to the artificial variable α in each additional constraint. The objective function f_i becomes a constraint function by changing the value of β_i from one to zero. For further details, see Nakayama et al. (1995). The idea of interchanging the roles of the functions is also handled in Subsections 5.10.3 and 5.12.7.

5.8.4. Implementation

A STOM implementation has been carried out in Bulgaria. The software is called MONP-16; see Vassilev et al. (1990). This program has been developed for nonlinear multiobjective optimization problems. The system suggests new aspiration levels for all the objective functions, and the user can freely change them. The feasibility of the aspiration levels is checked in the sense of linear approximations. If it is impossible to satisfy all the specified aspiration levels, the user can either modify the levels or go ahead and optimize anyway. In the latter case the solution will be weakly Pareto optimal but not satisficing. MONP-16 can handle objective functions both to be minimized and to be maximized.

5.8.5. Applications and Extensions

Some theoretical specifications concerning the STOM algorithm are presented in Nakayama and Furukawa (1985). The method is also applied to the aseismic design of a tower-pier system for a long span suspension bridge. Software implementing STOM for interactive construction accuracy control systems of cable-stayed bridges is introduced and applications are described in Nakayama et al. (1995, 1997). A linear diet problem and the erection of cable-stayed bridges are mentioned as STOM applications in Nakayama (1994). The latter problem is also handled in Nakayama (1995). An application to a water quality control problem of a river basin is presented in Nakayama (1985a) and Nakayama and Sawaragi (1984). In Olson (1993), the method is employed to solve a sausage blending problem and, in Nakayama et al. (1986), to solve a blending problem of industrial plastic materials. STOM is adapted for linear fractional objective functions in Nakayama (1991a) with an application concerning material blending in cement production. A blending problem in feed formulation for live stock is described in Nakayama (1995) as well as an interactive support system for bond trading. In addition, STOM is applied in a diet planning problem in Mitani and Nakayama (1997).

Sensitivity analysis of the STOM algorithm for linear problems is investigated in Nakayama (1989). Further, the relationship between STOM and fuzzy mathematical programming is handled in Nakayama (1995).

The ISWT method and STOM are combined in the method suggested in Wang (1992). The decision maker can choose the form of the interactive questions (surrogate worth values or new aspiration levels) at each iteration. The Pareto optimality of the solutions produced by the ε-constraint method is guaranteed by solving auxiliary problem (2.10.1) of Part I.

5.8.6. Concluding Remarks

STOM contains identical elements with STEM, the reference point method and the GUESS method. Therefore, the comments given there are not repeated here. The role of the decision maker is easy to understand. STOM requires even less input from the decision maker than the above-mentioned methods because only a part of the aspiration levels need to be given. The solutions obtained are properly Pareto optimal or weakly Pareto optimal depending on the scalarizing function used.

As said before, in practice, classifying the objective functions into three classes and specifying the amounts of increment and decrement for their values is a subset of specifying a new reference point. A new reference point is implicitly formed. Either the new aspiration levels are larger, smaller, or the same as in the current solution. Thus the same outcome can be obtained with different reasoning. A positive differentiating feature in STOM when compared to other classification-based methods is the automatic or exact trade-off. This decreases the amount of information inquired from the decision maker. STOM is in a sense opposite to STEM. In STOM, only desired improvements are specified, whereas only amounts of relaxation are used in STEM.

Because the method is based on satisficing decision making, the decision maker can freely search for a satisficing solution and change her or his mind, if necessary. No convergence based on value functions has even been intended.

5.9. Light Beam Search

Light beam search, described in Jaszkiewicz and Slowiński (1994, 1995), combines the reference point idea and tools of multiattribute decision analysis. That is why it is an interesting method for inclusion here to represent how the benefits of different problem solving areas can be put to use. Here we modify the original method for minimization problems.

5.9.1. Introduction

The basic setting in the light beam search is identical to the reference point method of Wierzbicki in the spirit of satisficing decision making. The achievement function to be minimized is function (3.5.3) where weighting coefficients are used only in the maximum part. They take into account the ideal and the nadir objective values. This achievement function means that ε-properly Pareto optimal solutions are generated. The reference point is here assumed to be an infeasible objective vector.

It is assumed that

1. Less is preferred to more by the decision maker.
2. The objective and the constraint functions are continuously differentiable.
3. The objective functions are bounded over the feasible region S.
4. None of the objective functions is more important than all the others together.

Assumption 3 is needed in order to have the ideal and the nadir objective vectors available. The other assumptions are related to the generation of alternative solutions.

In the light beam search it is acknowledged that reference points provide a practical and an easy way for the decision maker to direct the solution process. However, the learning process of the decision maker is supported better if the decision maker receives additional information about the Pareto optimal set at each iteration. This means that other solutions in the neighbourhood of the current solution (based on the reference point) are displayed. Thus far, the motivation is the same as in the reference point method. But what if the comparison of even a small number of alternative solutions is difficult for the decision maker? Or what if all the alternatives provided are indifferent to the decision maker? In such cases the decision maker may even stop the solution process and never get as far as the satisfactory solutions.

An attempt is made to avoid frustration on the part of the decision maker in the light beam search by the help of concepts used in multiattribute decision analysis and particularly in ELECTRE methods (see, for example, Roy (1990) and Vincke (1992, pp. 56–69)). The idea is to establish *outranking relations* between alternatives. It is said that the alternative $\mathbf{z}^1$ outranks the alternative $\mathbf{z}^2$, denoted by $\mathbf{z}^1 S\, \mathbf{z}^2$, if $\mathbf{z}^1$ is at least as good as $\mathbf{z}^2$. In the light beam search, additional alternatives near the current solution are generated so that they outrank the current one. Incomparable or indifferent alternatives are not shown to the decision maker.

To be able to compare alternatives and to define outranking relations, we need several thresholds from the decision maker. Assumption 4 is related to this. Because of the just noticeable difference or for some other reasons it is not always possible for the decision maker to distinguish between different alternatives. This means that there is an interval where indifference prevails.

For this reason the decision maker is asked to provide *indifference thresholds* q_i for each objective function ($i = 1, \ldots, k$). In fact the thresholds should be functions of the objective values, that is $q_i(z_i)$, but in the light beam search they are assumed to provide only local information and are thus constants.

The line between indifference and preference does not have to be sharp either. The hesitation between indifference and preference can be expressed by *preference thresholds* p_i for $i = 1, \ldots, k$. Applying the same reasoning as above, we assume here that p_i is not a function of the values of the objective function but constant. In addition, we must have $p_i \geq q_i \geq 0$ for $i = 1, \ldots, k$.

Given these thresholds we can distinguish three preference relations between pairs of alternative objective vectors ($\mathbf{z}^1$ and $\mathbf{z}^2$) for each component, that is, each objective function. We can say that as far as the ith components ($i = 1, \ldots, k$) of the two objective vectors are concerned,

$$\begin{aligned}
&\mathbf{z}^1 \text{ and } \mathbf{z}^2 \text{ are indifferent if} && |z_i^2 - z_i^1| \leq q_i \\
&\mathbf{z}^1 \text{ is weakly preferred to } \mathbf{z}^2 \text{ if} && q_i < z_i^2 - z_i^1 < p_i \\
&\mathbf{z}^1 \text{ is preferred to } \mathbf{z}^2 \text{ if} && z_i^2 - z_i^1 \geq p_i.
\end{aligned}$$

One more type of threshold, namely a *veto threshold* v_i for $i = 1, \ldots, k$ can be defined. It prevents a good performance in some components from compensating for poor values on some other components. As earlier, we assume the threshold to be constant and have the relation $v_i \geq p_i$ for $i = 1, \ldots, k$. In this case $\mathbf{z}^2$ cannot be preferred to $\mathbf{z}^1$ if $z_i^2 - z_i^1 \geq v_i$.

We can now define outranking relations on the basis of for how many components indifference, weak preference or preference is valid or preference cannot be valid. Let us compare the objective vector of the current iteration $\mathbf{z}^h$ and some other objective vector $\mathbf{z}$. Below, $\#i$ denotes the number of components, that is, objective functions, for which the condition mentioned holds. We define

$m_s(\mathbf{z}, \mathbf{z}^h)$ as $\#i$ where $\mathbf{z}$ is indifferent, weakly preferred or preferred to $\mathbf{z}^h$,
$m_q(\mathbf{z}^h, \mathbf{z})$ as $\#i$ where $\mathbf{z}^h$ is weakly preferred to $\mathbf{z}$,
$m_p(\mathbf{z}^h, \mathbf{z})$ as $\#i$ where $\mathbf{z}^h$ is preferred to $\mathbf{z}$,
$m_v(\mathbf{z}^h, \mathbf{z})$ as $\#i$ where $\mathbf{z}$ cannot be preferred to $\mathbf{z}^h$.

The outranking relations are defined according to the numbers above. If the decision maker has specified all the thresholds, that is the indifference, the preference and the veto thresholds, it is proposed in Jaszkiewicz and Slowiński (1994, 1995) that

$$\mathbf{z} S \mathbf{z}^h \text{ if } m_v(\mathbf{z}^h, \mathbf{z}) = 0,\ m_p(\mathbf{z}^h, \mathbf{z}) \leq 1 \text{ and } m_q(\mathbf{z}^h, \mathbf{z}) + m_p(\mathbf{z}^h, \mathbf{z}) \leq m_s(\mathbf{z}, \mathbf{z}^h)$$

be defined. This definition must be modified if no veto thresholds are available. In this case

$$\mathbf{z} S \mathbf{z}^h \text{ if } m_p(\mathbf{z}^h, \mathbf{z}) = 0 \text{ and } m_q(\mathbf{z}^h, \mathbf{z}) \leq m_s(\mathbf{z}, \mathbf{z}^h).$$

If no preference thresholds have been specified, the definition is

$$\mathbf{z} S \mathbf{z}^h \text{ if } m_v(\mathbf{z}^h, \mathbf{z}) = 0 \text{ and } m_q(\mathbf{z}^h, \mathbf{z}) \leq 1.$$

Finally, if only indifference thresholds are available, the outranking relation is defined by

$$\mathbf{z} S \mathbf{z}^h \text{ if } m_q(\mathbf{z}^h, \mathbf{z}) = 0.$$

5.9.2. Light Beam Algorithm

Let us now outline the light beam algorithm.

(1) If the decision maker wants to or can specify the best and the worst values for each objective function, denote the corresponding vectors by $\mathbf{z}^\star$ and $\mathbf{z}^{\text{nad}}$, respectively. Alternatively calculate $\mathbf{z}^\star$ and $\mathbf{z}^{\text{nad}}$. Set $h = 1$ and the reference point $\bar{\mathbf{z}}^h = \mathbf{z}^\star$. Initialize the set of saved solutions as $B = \emptyset$. Ask the decision maker to specify an indifference threshold for every objective function. If desired, (s)he can also specify preference and veto thresholds for them.

(2) Calculate a current solution $\mathbf{x}^h$ and the corresponding $\mathbf{z}^h$ by minimizing the achievement function with $\bar{\mathbf{z}}^h$.

(3) Present $\mathbf{z}^h$ to the decision maker. Calculate k Pareto optimal characteristic neighbours of $\mathbf{z}^h$ and present them as well to the decision maker. If the decision maker wants to see alternatives between any two of the $k + 1$ alternatives displayed, set their difference as a search direction, take different steps in this direction and project them onto the Pareto optimal set before showing them to the decision maker. If the decision maker wants to save the current solution $\mathbf{z}^h$, set $B = B \cup \{\mathbf{z}^h\}$.

(4) If desired, the decision maker can revise the thresholds. If this is the case, set $\mathbf{z}^h = \mathbf{z}^{h+1}$, $h = h + 1$ and go then to step (3). Otherwise, if the decision maker wants to give another reference point, denote it by $\bar{\mathbf{z}}^{h+1}$, set $h = h + 1$ and go to step (2). If, on the other hand, the decision maker wants to select one of the alternatives displayed or one solution in B as a current solution, set it as $\mathbf{z}^{h+1}$, set $h = h + 1$ and go to step (3). Finally, if one of the alternatives is satisfactory, set the corresponding decision vector to be $\mathbf{x}^{h+1}$, set $h = h + 1$ and go to step (5).

(5) Stop with $\mathbf{x}^h$ as the final solution.

The option of saving desirable solutions in set B increases the flexibility of the method. The decision maker can explore different directions and select the best among different trials. The possibility of having a look at solutions between any two alternatives is related to the GDF method. The same idea will also be handled in Section 5.12. The alternative solutions in step (3) can be projected by minimizing the achievement function.

Let us consider how the Pareto optimal characteristic neighbours of $\mathbf{z}^h$ are generated. The thresholds specified by the decision maker are needed in

defining outranking relations as described in Subsection 5.9.1. *Characteristic neighbours* are new alternative objective functions that outrank the current solution. The number of characteristic neighbours $\mathbf{z}(i)$ is equal to the number of objective functions. For each $i = 1, \ldots, k$, the neighbour $\mathbf{z}(i)$ is the point in the outranking neighbourhood of $\mathbf{z}^h$ with maximal distance from $\mathbf{z}^h$ in the direction where the value of the ith component locally improves most.

The neighbours are determined by projecting the gradient of one objective function at a time onto the linear approximation of those constraints that are active in $\mathbf{z}^h$ with gradient projection methods (this necessitates differentiability and, thus, assumption 2); see Jaszkiewicz and Slowiński (1994, 1995) for details. The feasible direction in the objective space offering the greatest improvement for the ith component of $\mathbf{z}^h$ is denoted by $\mathbf{d}^i$. The outranking characteristic neighbour in that direction is obtained with the problem

$$\begin{array}{lll} & \text{maximize} & \alpha \\ (5.9.1) & \text{subject to} & \mathbf{z}(i) S\, \mathbf{z}^h \quad \text{where} \quad \mathbf{z}(i) = \mathbf{z}^h + \alpha \mathbf{d}^i, \\ & & \alpha \geq 0. \end{array}$$

After solving problem (5.9.1) for each $i = 1, \ldots, k$, we have k characteristic neighbours. Each $\mathbf{z}(i)$ is projected onto the Pareto optimal set before being displayed to the decision maker. This can be carried out, for example, by minimizing the achievement function with each neighbour as a reference point.

5.9.3. Comments

The idea of the light beam search is analogous to projecting a focused beam of light from the reference point onto the Pareto optimal set. The lighted part of the Pareto optimal set changes if the location of the spotlight, that is, the reference point or the point of interest in the Pareto optimal set are changed. This connection explains the name of the method. An implementation of the light beam search is available from its developers (see Section 2.2 in Part III).

The light beam search can be characterized as an ad hoc method. If a value function were available, it could not directly determine new reference points. It could, however, be used in comparing the set of alternatives. Yet, the thresholds are important in the method and they must come from the decision maker.

This method combines elements of multiobjective optimization and multiattribute decision analysis in an interesting way. An extension is suggested in Wierzbicki (1997b), where both aspiration levels forming a reference point and reservation levels (to be avoided) are used. In this case the reference point still determines the source of light but the reservation levels are used to generate a cone of light. Some convergence ideas are put forward in Wierzbicki (1997b) as well.

5.9.4. Concluding Remarks

The light beam search is a rather versatile solution method where the decision maker can specify reference points, compare a set of alternatives and affect the set of alternatives in different ways. Thresholds are used to try to make sure that the alternatives generated are not worse than the current solution. In addition, they are different enough to be compared and comparable on the whole. This should decrease the burden on the decision maker.

Specifying different thresholds is a new aspect when compared to the methods presented earlier. This may be demanding for the decision maker. Anyway, it is positive that the thresholds are not assumed to be global but can be altered at any time. In other words, outranking relations based on the threshold values are only used as local preference models in the neighbourhood of the current solution.

The idea of combining strengths from different areas certainly deserves further study. Nevertheless, this approach also has its weaknesses. As noted in Jaszkiewicz and Słowiński (1994, 1995), it may be computationally rather demanding to find the exact characteristic neighbours in a general case. Parallel computing is one solution. If this is not possible, one can at least present different neighbours as soon as they are calculated instead of waiting till all of them have been generated. The visualization of alternatives is handled in Chapter 3 of Part III.

5.10. Reference Direction Approach

The reference direction approach was introduced in Korhonen and Laakso (1984, 1985, 1986a) by the name visual interactive approach. It contains ideas from, for example, the GDF method and the reference point method of Wierzbicki. However, more information is provided to the decision maker. The algorithm works best for MOLP problems if it is desired to check the optimality of the final solution. Otherwise, the algorithm can be applied to nonlinear problems as well. The algorithm was originally designed for the maximization of problems but here it is presented in the form of minimization. The reference direction approach and its extensions are also briefly described in Korhonen (1997).

5.10.1. Introduction

In reference point-based methods, a reference point consisting of aspiration levels for each objective function is projected onto the Pareto optimal set by an achievement function. This idea is extended here so that a whole so-called reference direction is projected onto the Pareto optimal set. The *reference direction* is a vector from the current iteration point to the reference point. After

the projection the decision maker can examine this Pareto optimal curve or a representation of it by the means of computer graphics.

An interesting feature in the reference direction approach is that no explicit knowledge is assumed about the properties of the value function during the solution process. However, sufficient conditions for optimality can be established for the termination point of the algorithm, if the decision maker's underlying value function is assumed to be pseudoconcave (and differentiable) at that point (and several other assumptions to be listed later are fulfilled). The optimality conditions are necessary only for MOLP problems.

5.10.2. Reference Direction Approach Algorithm

The algorithm is as follows. Once again, in the notation we employ objective vectors for simplicity. Naturally, the actual calculations are performed in the decision variable space.

(1) Find an arbitrary starting objective vector $\mathbf{z}^1 \in \mathbf{R}^k$. Set $h = 1$.

(2) Ask the decision maker to specify a reference point $\bar{\mathbf{z}}^h \in \mathbf{R}^k$ and set $\mathbf{d}^{h+1} = \bar{\mathbf{z}}^h - \mathbf{z}^h$ as a new reference direction.

(3) Find the set Z^{h+1} of (weakly) Pareto optimal solutions $\mathbf{z}$ that solve the problem

$$\begin{aligned} &\text{minimize} && s_{\bar{\mathbf{z}},\mathbf{w}}(\mathbf{z}) \\ &\text{subject to} && \bar{\mathbf{z}} = \mathbf{z}^h + t\mathbf{d}^{h+1}, \\ & && \mathbf{z} \in Z \text{ is Pareto optimal,} \end{aligned}$$

where $s_{\bar{\mathbf{z}},\mathbf{w}}$ is an achievement function, $\mathbf{w}$ is a weighting vector and t has different discrete nonnegative values.

(4) Ask the decision maker to select the most preferred solution $\mathbf{z}^{h+1}$ in Z^{h+1}.

(5) If $\mathbf{z}^h \neq \mathbf{z}^{h+1}$, set $h = h + 1$ and go to step (2). Otherwise, check the optimality conditions. If the conditions are satisfied, stop with $\mathbf{x}^{h+1}$ corresponding to $\mathbf{z}^{h+1}$ as the final solution. Otherwise, set $h = h + 1$ and set $\mathbf{d}^{h+1}$ to be a new search direction identified by the optimality checking procedure. Go to step (3).

The setting of the algorithm makes it possible for the starting point to be any point in the objective space. It does not have to be feasible, much less Pareto optimal, since it is projected onto the (weakly) Pareto optimal set in step (3). As a weighting vector one can use, for example, the reference point specified by the decision maker.

The straight line from the current iteration point $\mathbf{z}^h$ (or its Pareto optimal projection at the first iteration) to the boundary of the Pareto optimal set is discretized and projected onto the set of Pareto optimal points. The discretization means using several different values for t. For linear problems parametric linear programming can be used to obtain a Pareto optimal curve when the parameter t has values from zero to infinity. The idea is to plot the obtained values

of the objective functions on a computer screen as value paths (see Subsection 3.3.1 in Part III) with different colours illustrating each of the objectives. The decision maker can move the cursor back and forth and see the corresponding numerical values at each point.

The achievement function $s_{\bar{\mathbf{z}},\mathbf{w}}(\mathbf{z})$ is of the same form as presented in Section 3.5, namely

$$s_{\bar{\mathbf{z}},\mathbf{w}}(\mathbf{z}) = \max_{i \in I} \frac{z_i - \bar{z}_i}{w_i}, \tag{5.10.1}$$

where $I = \{i \mid w_i > 0\} \subset \{1, \ldots, k\}$, $\mathbf{w}$ is a weighting vector, $\mathbf{z} \in Z$ is an objective vector to be searched for and $\bar{\mathbf{z}} \in \mathbf{R}^k$ is a reference point. Note that if it is desired to avoid weakly Pareto optimal solutions, then an augmentation term can be added to the achievement function as, for example, in Subsection 3.4.5 or Section 3.5 (see also Steuer (1986, pp. 422–431)). An alternative is suggested in Korhonen (1997) and Korhonen and Halme (1996), where lexicographic ordering is used to guarantee the Pareto optimality of the solutions. However, originally such kind of actions were not considered to be necessary because the purpose was simply to produce different solutions effectively. Distance measure (5.10.1) has been chosen to facilitate parametric linear programming (even though the solutions are only guaranteed to be weakly Pareto optimal).

The minimization problem of $s_{\bar{\mathbf{z}},\mathbf{w}}$ is nondifferentiable but it can be transformed into an equivalent, differentiable form assuming the differentiability of the functions involved. Let us, for clarity, formulate the problem in the decision variable space where it is solved as

$$\begin{aligned} &\text{minimize} && \alpha \\ &\text{subject to} && f_i(\mathbf{x}) - \alpha w_i \leq z_i^h + t d_i^{h+1} \quad \text{for all} \quad i \in I, \\ & && \mathbf{x} \in S, \end{aligned}$$

with $\mathbf{x} \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$ as variables.

Checking the optimality conditions in step (5) is the most complicated part of the algorithm. Thus far, no specific assumptions have been set on the value function. It may change during the solution process or it may not even exist at all. It is only assumed that

1. Less is preferred to more by the decision maker.

However, we can check whether a given objective vector $\mathbf{z}^{h+1}$ (and the corresponding decision variable $\mathbf{x}^{h+1}$) is optimal for the value function, assuming that

2. The underlying value function $U\colon \mathbf{R}^k \to \mathbf{R}$ exists and is pseudoconcave on Z.
3. The feasible region S is convex and compact.
4. The constraint functions are differentiable.

Let the feasible directions at $\mathbf{z}^{h+1}$ be denoted by $\mathbf{d}(j)$, $j = 1, \ldots, p$. We define a cone C containing all those feasible directions by

$$C = \left\{ \mathbf{z} \in Z \mid \mathbf{z} = \mathbf{z}^{h+1} + \sum_{j=1}^{p} \beta_j \mathbf{d}(j),\ \beta_j \geq 0 \right\}. \tag{5.10.2}$$

If we have $\mathbf{z}^h = \mathbf{z}^{h+1}$ in step (5), we know that the projection of $\mathbf{d}^h$ in the weakly Pareto optimal set is not a direction of improvement. Then we can apply the following sufficient condition for optimality.

Theorem 5.10.1. Let assumptions 2–4 be satisfied. Let $\mathbf{z}^{h+1} \in Z$ and let C be a cone containing all the feasible directions at $\mathbf{z}^{h+1}$ (as in (5.10.2)). Let us assume that

$$U(\mathbf{z}^{h+1}) \geq U(\mathbf{z}^{h+1} + \beta_j \mathbf{d}(j)) \quad \text{for all} \quad \beta_j \geq 0 \text{ and } j = 1, \ldots, p.$$

Then $\mathbf{z}^{h+1}$ is a globally optimal solution (with respect to U).

Proof. See Korhonen and Laakso (1986a).

For MOLP problems we know that if the current solution is not optimal, then one of the feasible directions of cone C must be a direction of improvement. This direction is then used as a new reference direction in step (3). In other words, to be able to apply Theorem 5.10.1 at a certain point, the decision maker must first check every feasible direction at that point for improvement. This increases both the computational costs and the burden on the decision maker. It is demonstrated in Halme and Korhonen (1989) and Korhonen and Laakso (1986a, b) how the number of search directions can be reduced. For nonlinear problems the cone containing all the feasible directions may consist of an infinite number of generators. In this case, the optimality cannot be checked in practice (an infinite number of checks would be needed).

5.10.3. Comments

Note that the termination condition of Theorem 5.10.1 is analogous to the Karush-Kuhn-Tucker optimality conditions. This is proved in Halme and Korhonen (1989). If the value function is known, it is easy to compare alternative objective vectors. However, what was said concerning the difficulty in determining new reference points in connection with the reference point method in Section 5.6 is also valid here. Thus the reference direction approach can be characterized as an ad hoc method.

The graphical illustration of the alternatives has been an important aspect in the development of multiobjective optimization methods that seek to improve and facilitate the co-operation between the decision maker and the analyst (computer). That is why graphical illustration of the alternatives is here emphasized.

The computation time for large problems can be reduced by presenting one piece at a time of the weakly Pareto optimal curve or its representation to the

decision maker. If (s)he finds the end point to be the most satisfactory one, then the next piece can be presented. If the number of objective functions is large, the quality of graphical illustration suffers. For this reason, it is advisable not to have more than ten objective functions at a time.

If it is not desired to check the optimality of the final result, the problem to be solved does not have to satisfy any special assumptions. This means that the reference direction approach can be applied to more general problems. The reverse is valid as well. If the assumptions set are not satisfied, the optimality cannot be checked, but the method can, of course, be used in any other way.

A similar interactive line search algorithm for MOLP problems is presented in Benson and Aksoy (1991). The procedure generates only Pareto optimal points and is able to automatically correct possible errors in the decision maker's judgement.

The ideas of the reference direction approach are adapted to the goal programming environment in Korhonen and Laakso (1986b). The intention is to relax the predetermined roles of the objective functions and the constraints, that is, to enable the roles to be interchanged. For that reason, the problem to be solved is now assumed to be in the generalized goal programming form (see Section 4.3). The objective functions are considered to be flexible goals and the constraint functions inflexible goals. At each iteration, the decision maker can easily convert flexible goals into inflexible ones and vice versa. This increases the freedom of the decision maker. Combining achievement functions into goal programming also eliminates the problems caused by feasible aspiration levels (see Section 4.3).

The idea of changing the roles of the functions is refined in Korhonen and Narula (1993). A systematic way of changing the roles of the objective functions and the constraints is described therein. The presentation examines where and how the changes can be carried out. This systematic handling concerns MOLP problems, but the idea can in principle be generalized to other problems.

A dynamic user interface to the reference direction approach and its adaptation to generalized goal programming is introduced in Korhonen and Wallenius (1988). This method has been designed for MOLP problems and is called the *Pareto race*. The software system implementing the Pareto race is called VIG (Visual Interactive Goal programming) and it is described in Korhonen (1987, 1990, 1991a) and Korhonen and Wallenius (1989c, 1990). VIG is a dynamic, visual and interactive solution system for MOLP problems with the emphasis on graphical illustration.

The Pareto race develops reference directions in a dynamic way. In VIG, the reference directions and the step-sizes are updated according to the actions of the decision maker who can thus feel that (s)he is in control. The decision maker can travel around the (weakly) Pareto optimal set as if driving a car. The pioneering ideas of realizing user interfaces in VIG are supported by a comparison of five MOLP programs in Korhonen and Wallenius (1989b). VIG

was found to be superior. The main reason was that the decision makers found the aspiration levels to be a comfortable way of expressing preference relations.

The Pareto race is extended into a computer graphics-based decision support system in Korhonen et al. (1992b). The new method is especially useful for large-scale MOLP problems.

5.10.4. Concluding Remarks

In the reference direction approach the role of the decision maker is reminiscent of the reference point method. (S)he has to both specify reference points and select the most preferred alternatives. In the reference point methods, however, there are fewer choices to select from. If the problem is set in a generalized goal programming form, the decision maker can also interchange the roles of the objective and the constraint functions. By the reference direction approach, the decision maker can explore a wider part of the weakly Pareto optimal set than by the reference point method, even by providing similar reference point information. This possibility brings the task of comparing the alternatives and selecting the most preferred of them.

The reference direction approach works best for MOLP problems, as it has basically been designed for them. It is interesting that the method requires no additional assumptions about the problem and the underlying value function until the optimality of the final solution is to be examined. The optimality can be guaranteed under certain assumptions and with some effort.

The performance of the method depends greatly on how well the decision maker manages to specify the reference directions that lead to improved solutions. Korhonen and Laakso (1986a) mention that particularly when the number of objective functions is large, the specification of reference points may be quite laborious for the decision maker. In this case, they suggest that random directions in conjunction with decision maker-defined reference directions should be used. See Korhonen and Laakso (1986a) for a discussion concerning other ways of specifying the reference directions. Naturally, the choice of the weighting coefficients affects the direction of the projection even though the selection of their values has not been stressed here.

The consistency of the decision maker's answers is not important and it is not checked in the algorithm. Thus the algorithm may cycle. This can also be seen as a positive feature, since the decision maker is able to return to such parts that (s)he already has examined, if (s)he changes her or his mind.

5.11. Reference Direction Method

The reference direction (RD) method, introduced in Narula et al. (1994a, b) is closely related to the reference direction approach. As its name suggests it is also based on reference directions. To avoid confusion between these two methods with very similar names we use the name RD method when referring to the reference direction method in what follows. The RD method has been designed for nonlinear maximization problems but here we revise it for minimization and generalize it. In addition, we relax the original convexity assumption and settle for local optima.

5.11.1. Introduction

In the RD method, objective function values $\mathbf{z}^h$ calculated at a point $\mathbf{x}^h$ are presented to the decision maker and (s)he is asked to specify a reference point $\bar{\mathbf{z}}^h$ consisting of desired levels for the objective functions. Once again, we move around the weakly Pareto optimal set, which is why some objective functions must be allowed to increase in order to attain lower values for some other objective functions. In other words, some components of the reference point have to be lower and some others higher or equal when compared to the current solution. Allowing the set of higher values to be empty is a generalization of the original form of the method. (Weakly Pareto optimal solutions can be made Pareto optimal.)

As mentioned earlier, specifying a reference point is equivalent to an implicit classification using three classes and indicating those objective functions whose values should be decreased till they reach some acceptable aspiration level, those whose values are satisfactory at the moment, and those whose values are allowed to increase to some upper bound. Let us denote the sets of functions by $I^<$, $I^=$ and $I^>$, respectively. We denote the components of the reference point corresponding to the set $I^>$ by ε_i^h because we have upper bounds in question. To put it briefly, a reference point is here sensible and the corresponding classification feasible if $I^< \neq \emptyset$ and $I^> \cup I^= \neq \emptyset$.

It is once again assumed that

1. Less is preferred to more by the decision maker.

The reference direction $\bar{\mathbf{z}}^h - \mathbf{z}^h$ is a fundamental element in the RD method. The decision maker specifies a priori the number of steps to be taken in the reference direction. The idea is to move step by step as long as the decision maker wants to. In other words, extra computation is avoided by calculating only those alternatives the decision maker wants to see.

The alternatives are produced by solving the *RD problem*

$$
\begin{aligned}
& \text{minimize} && \max_{i \in I^<} \frac{f_i(\mathbf{x}) - z_i^h}{z_i^h - \bar{z}_i^h} \\
& \text{subject to} && f_i(\mathbf{x}) \leq \varepsilon_i^h + \alpha(z_i^h - \varepsilon_i^h) \quad \text{for all} \quad i \in I^>, \\
& && f_i(\mathbf{x}) \leq z_i^h \quad \text{for all} \quad i \in I^=, \\
& && \mathbf{x} \in S,
\end{aligned} \tag{5.11.1}
$$

where $\mathbf{z}^h$ is the current solution, $0 \leq \alpha < 1$ is the step-size in the reference direction, $\bar{z}_i^h < z_i^h$ for $i \in I^<$ and $\varepsilon_i^h > z_i^h$ for $i \in I^>$. The problem is nondifferentiable but it can be transformed into a differentiable form by introducing an additional variable as described earlier (see, e.g., problem (3.4.3)). If some of the objective or the constraint functions are nondifferentiable, a single objective solver applicable to nondifferentiable problems is needed.

The RD problem produces weakly Pareto optimal solutions.

Theorem 5.11.1. The solution of RD problem (5.11.1) is weakly Pareto optimal for every $0 \leq \alpha < 1$.

Proof. Let $\mathbf{x}^* \in S$ be a solution of the RD problem for some $0 \leq \alpha < 1$. Let us assume that it is not weakly Pareto optimal. In this case there exists some point $\mathbf{x}^\circ \in S$ such that $f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*)$ for every $i = 1, \ldots, k$.

Because $\mathbf{x}^*$ is feasible in problem (5.11.1), $\mathbf{x}^\circ$, being weakly Pareto optimal, must also be feasible. In addition, $z_i^h - \bar{z}_i^h > 0$ for every $i \in I^<$ and that is why

$$
\frac{f_i(\mathbf{x}^\circ) - z_i^h}{z_i^h - \bar{z}_i^h} < \frac{f_i(\mathbf{x}^*) - z_i^h}{z_i^h - \bar{z}_i^h} \quad \text{for every} \quad i \in I^<.
$$

This implies that

$$
\max_{i \in I^<} \frac{f_i(\mathbf{x}^\circ) - z_i^h}{z_i^h - \bar{z}_i^h} < \max_{i \in I^<} \frac{f_i(\mathbf{x}^*) - z_i^h}{z_i^h - \bar{z}_i^h}
$$

and, thus, $\mathbf{x}^*$ cannot be a solution of problem (5.11.1). This contradiction completes the proof and $\mathbf{x}^*$ is weakly Pareto optimal. □

A result concerning the opposite direction and Pareto optimality can also be established.

Theorem 5.11.2. Let $\mathbf{x}^* \in S$ be Pareto optimal. Then there exists a reference point and a real number $0 \leq \alpha < 1$ such that $\mathbf{x}^*$ is a solution of RD problem (5.11.1).

Proof. Let $\mathbf{x}^* \in S$ be Pareto optimal. Let us assume that there does not exist $\bar{\mathbf{z}}$ and α such that $\mathbf{x}^*$ is a solution of the RD problem. Let us suppose that

we have the current solution of the RD problem $\mathbf{x}^h$ and the corresponding $\mathbf{z}^h$ available.

Let us choose $\mathbf{f}(\mathbf{x}^*)$ as a reference point. This means that we set $\bar{z}_i = f_i(\mathbf{x}^*)$ for those indices where $z_i^h > f_i(\mathbf{x}^*)$ and we denote this index set by $I^<$. Further, we set $\varepsilon_i = f_i(\mathbf{x}^*)$ for indices $i \in I^>$ satisfying $z_i^h < f_i(\mathbf{x}^*)$. Finally, the set of indices where $z_i^h = f_i(\mathbf{x}^*)$ is valid is denoted by $I^=$. This setting is possible because $\mathbf{x}^*$ is assumed to be Pareto optimal and $\mathbf{x}^h$ is weakly Pareto optimal according to Theorem 5.11.1. That is why $I^< \neq \emptyset$ and $I^= \cup I^> \neq \emptyset$. In addition we set $\alpha = 0$.

Because $\mathbf{x}^*$ is not a solution of the RD problem, there exists another point $\mathbf{x}^\circ \in S$ that is a solution of the RD problem, meaning that

$$\max_{i \in I^<} \frac{f_i(\mathbf{x}^\circ) - z_i^h}{z_i^h - f_i(\mathbf{x}^*)} < \max_{i \in I^<} \frac{f_i(\mathbf{x}^*) - z_i^h}{z_i^h - f_i(\mathbf{x}^*)} = -1.$$

Thus $f_i(\mathbf{x}^\circ) - z_i^h < -(z_i^h - f_i(\mathbf{x}^*))$, that is, $f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*)$ for all $i \in I^<$. Because $\mathbf{x}^\circ$ is a solution of problem (5.11.1), it must be feasible. In other words, we have $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*) + 0$ for $i \in I^>$ and $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$ for $i \in I^=$. Here we have a contradiction to the assumption that $\mathbf{x}^*$ is Pareto optimal. This completes the proof and $\mathbf{x}^*$ must be a solution of the RD problem. □

According to Theorem 5.11.2 we know that any Pareto optimal solution can be found with an appropriate classification.

An augmented formulation of the RD problem is presented in Narula et al. (1994a, b) in order to produce only Pareto optimal solutions.

5.11.2. RD Algorithm

The steps of the RD algorithm are the following:

(1) Calculate a starting solution $\mathbf{x}^1$ by solving auxiliary problem (5.11.2). Show the corresponding objective vector $\mathbf{z}^1$ to the decision maker. If (s)he wants to stop, go to step (5). Otherwise, set $h = 1$.

(2) If the decision maker does not want to decrease any component of $\mathbf{z}^h$, go to step (5). Otherwise, ask the decision maker to specify a reference point $\bar{\mathbf{z}}^h$, where some of the components are lower and some higher or equal when compared to those of $\mathbf{z}^h$. If there are no higher values, set $P = r = 1$ and go to step (3). Otherwise ask the decision maker also to specify the maximum number of alternatives P (s)he wants to see. Set $r = 1$.

(3) Set $\alpha = 1 - r/P$. Solve RD problem (5.11.1) to obtain a solution $\mathbf{x}^h(r)$ and the corresponding $\mathbf{z}^h(r)$. Set $r = r + 1$.

(4) Show $\mathbf{z}^h(r)$ to the decision maker. If (s)he is satisfied, go to step (5). If $r \leq P$ and the decision maker wants to see another solution, go to step (3). Otherwise, if $r > P$ or the decision maker wants to change the reference point, set $\mathbf{z}^{h+1} = \mathbf{z}^h(r)$, $h = h + 1$ and go to step (2).

(5) Stop with $\mathbf{x}^h$ corresponding to $\mathbf{z}^h$ as the final solution.

The starting solution is calculated by solving the problem

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} f_i(\mathbf{x}) \\ &\text{subject to} && \mathbf{x} \in S. \end{aligned} \tag{5.11.2}$$

Naturally, this problem can be formulated as a differentiable problem with, if necessary and possible, the help of an additional variable (as earlier).

5.11.3. Comments

The RD method is an ad hoc method. The existence of a value function would not help in specifying reference points or the numbers of steps to be taken. It could not even help in selecting the most preferred alternative. The reason is that one must decide for one point at a time whether to calculate new alternatives or not. If the new alternative turned out to be less preferred than its predecessor, one could not go back anyway.

In Miettinen and Mäkelä (1998a), a water quality management problem is solved by the RD method. A modification of the RD method for convex, nonlinear integer problems is introduced in Gouljashki et al. (1997).

5.11.4. Concluding Remarks

The RD method can be considered an interactive classification-based method. It does not require artificial or complicated information from the decision maker; only reference points and the number of intermediate solutions are used. The decision maker is not asked to compare several different alternatives but only to decide whether another alternative is to be generated or not.

The decision maker must a priori determine the number of steps to be taken in the reference direction, and then intermediate solutions are calculated one by one as long as the decision maker wants to. This can be seen as a benefit as well as a weakness. On the one hand, it is computationally efficient since it may be unnecessary to calculate all the intermediate solutions. On the other hand, the decision maker is unable to return to a solution once it has been discarded, which may be a disadvantage. Further, the number of steps to be taken cannot be changed.

Thus far, we have described several different methods for multiobjective optimization. The question of differentiability has not been emphasized. However, nondifferentiability and many kinds of irregularities and discontinuities are characteristic of real-world optimization problems, for example, in economics and engineering. For this reason, it is important to have methods that are able to solve nondifferentiable problems. We shall now handle nondifferentiability aspects. Remember that by nondifferentiability we mean that all the objective and the constraint functions are locally Lipschitzian.

One way of handling nondifferentiabilities is to regularize them into a differentiable form by utilizing some smoothing techniques. However, regularization simplifies the problem and causes errors in the model. Depending on the smoothing parameters, the regularized problem is numerically either unreliable or unstable. That is why it is important to be able to solve nondifferentiable problems as they are, without simplifications.

For example, the area of optimal control has complex problems of a multiobjective nature containing nondifferentiable functions. Traditionally, they have been solved (e.g., in Haslinger and Neittaanmäki (1988)) by first scalarizing the multiple objective functions into one by some simple method (like the weighting method) and then regularizing the nondifferentiabilities. After discretization, the problems have been solved by traditional, differentiable single objective optimization methods. The drawbacks of regularization were mentioned above. In scalarization, the appropriate weighting coefficients are difficult to specify. If some of the objective functions originate in technological constraints, the weighting method may bring about inaccuracies and the solution may be irrelevant in a technological sense. For this reason, it is important to use interactive methods, where the user can direct the solution process in a desirable direction.

All the noninteractive methods presented in Chapters 2 to 4 can be employed with nondifferentiable problems whenever the single objective solvers utilized can handle nondifferentiable functions. It is to be noted that the MPB method in Section 2.2 has particularly been designed for nondifferentiable problems.

The case is different with interactive methods. Many of the interactive methods described thus far assume that the functions involved are differentiable. This is true especially for methods based on maximizing the underlying value function. Most classification- and reference point-based methods do not guide the decision maker, nor do they assume differentiability. Of the interactive methods described, the Tchebycheff method, STEM, the reference point method, the GUESS method and the RD method can all be used to solve nondifferentiable problems assuming that a nondifferentiable single objective solver is available.

Few methods especially designed to handle nondifferentiable problems have been proposed. In the sequel, we present one such method, known as NIMBUS. Another aim in developing NIMBUS has been in trying to overcome some of

the weaknesses detected in the older methods. Most of the methods previously described have had an effect on the development of NIMBUS. Either they have offered useful ideas to adopt or unsatisfactory properties to avoid.

Trade-off rate information cannot be exploited in nondifferentiable problems in the way it is used in the ISWT method and in SPOT and STOM. The natural reason is that obtaining trade-off information from the Karush-Kuhn-Tucker multipliers necessitates that the functions are twice continuous differentiable. How to obtain trade-off information in nondifferentiable cases needs and deserves more research.

The ideas of reference points and satisficing decision making seem to be generalizable to nondifferentiable problems. We can adopt the ideas of classifying the objective functions and reference points and mix them with some ideas from nondifferentiable analysis. The outcome is described in the next section.

5.12. NIMBUS Method

NIMBUS (Nondifferentiable Interactive Multiobjective BUndle-based optimization System), presented in Miettinen (1994) and Miettinen and Mäkelä (1995, 1997), is an interactive multiobjective optimization method designed especially to be able to handle nondifferentiable functions efficiently. For this reason, it is capable of solving complicated real-world problems. We introduce two versions of NIMBUS. They have differences in both their theoretical and computational aspects. Theoretically, the versions differ in handling the information requested from the user. Numerical experiments indicate differences in the computational efficiency and controllability of the solution processes.

5.12.1. Introduction

The starting point in developing the NIMBUS method has been somewhat the opposite to theoretical soundness. Emphasizing theoretical aspects may lead to difficulties on the decision maker's side and more or less instable results, not to mention higher computational costs. In the NIMBUS method, the idea has been to overcome the difficulties encountered with many other interactive methods. The most important aspects have appeared to be the effectiveness and the comfortableness of the decision maker. Thus, the interaction phase has been aimed at being comparatively simple and easy to understand for the decision maker. NIMBUS offers flexible ways of performing interactive evaluation of the problem and determining the preferences of the decision maker during the solution process. At each iteration of the interactive solution process the decision maker can direct the search according to her or his wishes.

Aspiration levels and classification have been selected as the means of interaction between the decision maker and the algorithm. It has been emphasized on several occasions (e.g., in Nakayama (1995)) that an aspiration level-based

approach is effective in practical fields. Among the validating facts for this statement are the following. Aspiration levels do not require consistency from the decision maker and they reflect her or his wishes well. In addition, they are easy to implement. Using aspiration levels as a way of receiving information from the decision maker means avoiding difficult and artificial concepts.

It is assumed that

1. Less is preferred to more by the decision maker.
2. The objective and the constraint functions are locally Lipschitzian.
3. The objective functions are bounded (from below) over the feasible region S.

The second assumption comes from nondifferentiable analysis, and the third assumption from the requirement of having the ideal objective vector available.

In the classification of the objective functions, the decision maker can easily indicate what kind of improvements are desirable and what kind of impairments are tolerable. The idea is that the decision maker examines at every iteration h the values of the objective functions calculated at the current solution $\mathbf{x}^h$ and divides the objective functions into up to five classes. The classes are functions f_i whose values

- should be decreased ($i \in I^<$),
- should be decreased to a certain aspiration level $\bar{z}_i^h$ ($i \in I^\leq$),
- are satisfactory at the moment ($i \in I^=$),
- are allowed to increase to a certain upper bound ε_i^h ($i \in I^>$), and
- are allowed to change freely ($i \in I^\diamond$),

where $I^< \cup I^\leq \cup I^= \cup I^> \cup I^\diamond = \{1, \ldots, k\}$, $I^< \cup I^\leq \neq \emptyset$ and $I^= \cup I^> \cup I^\diamond \neq \emptyset$.

In addition to the classification, the decision maker is asked to specify the aspiration levels $\bar{z}_i^h$ for $i \in I^\leq$ satisfying $\bar{z}_i^h < f_i(\mathbf{x}^h)$ and the upper bounds ε_i^h for $i \in I^>$ such that $\varepsilon_i^h > f_i(\mathbf{x}^h)$. Notice that the two somewhat parallel classes $I^<$ and $I^\leq$ are available. The difference between them is that the functions in $I^<$ are to be minimized as far as possible but the functions in $I^\leq$ only as far as the aspiration level. Thus the functions in the latter class are called *aspiration functions*.

The classification is the core of NIMBUS. However, the decision maker can tune the order of importance inside classes $I^<$ and $I^\leq$ with optional positive weighting coefficients w_i^h summing up to one (for numerical stability). If the decision maker does not want to specify any weighting coefficients, they are set equal to one. Note that the weighting coefficients do not change the primary orientation specified in the classification phase.

NIMBUS has more classes than STEM, STOM or the RD method. In this way the decision maker has more freedom in specifying the desired changes in the objective values and (s)he can select a class reflecting her or his desires best. The class $I^\diamond$ is new when compared to the methods described thus far. In practice, it means that not all the objective functions have to be classified at all. Naturally all the classes do not have to be employed.

After the decision maker has classified the objective functions, one of the two alternative subproblems, called vector and scalar subproblems, is formed. Thus, the original multiobjective optimization problem is transformed into either a new multiobjective or a single objective optimization problem, accordingly. The subproblems lead to two different versions of NIMBUS, to be called *vector version* and *scalar version*. We first introduce the older, that is, the vector version.

5.12.2. Vector Subproblem

According to the classification and the connected information, a *vector subproblem*

$$\begin{aligned} & \text{minimize} && \left\{ f_i(\mathbf{x}) \ (i \in I^<),\ \max_{j \in I^\leq} \left[\max \left[f_j(\mathbf{x}) - \bar{z}_j^h,\ 0 \right] \right] \right\} \\ & \text{subject to} && f_i(\mathbf{x}) \leq f_i(\mathbf{x}^h) \ \text{ for all } \ i \in I^=, \\ & && f_i(\mathbf{x}) \leq \varepsilon_i^h \ \text{ for all } \ i \in I^>, \\ & && \mathbf{x} \in S \end{aligned} \tag{5.12.1}$$

is formed (see Miettinen (1994) and Miettinen and Mäkelä (1995)).

The vector subproblem seems to be even more complicated than the original problem. Nonetheless, the advantage of this formulation is the fact that the opinions and the hopes of the decision maker are taken carefully into account. Notice that if $I^\leq \neq \emptyset$, we have a nondifferentiable problem regardless of the differentiability of the original problem. This fact does not bring any additional difficulties since we are in any case prepared for handling nondifferentiabilities.

The vector version is quite general. The classification of the objective functions can be performed as if the ε-constraint method, the weighting method, lexicographic ordering or goal programming were used to produce new solutions.

In order to be able to solve the vector subproblem, we need the MPB method (introduced in Section 2.2). If the constraints are inequalities, that is $S = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T \leq \mathbf{0}\}$, the improvement function $H\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ applied to problem (5.12.1) is of the form

$$\begin{aligned} H(\mathbf{x}^1, \mathbf{x}^2) = \max \Big\{ & f_i(\mathbf{x}^1)/w_i^h - f_i(\mathbf{x}^2)/w_i^h,\ (i \in I^<), \\ & \max_{j \in I^\leq} \left[\max \left[f_j(\mathbf{x}^1)/w_j^h - \bar{z}_j^h,\ 0 \right] \right] - \max_{j \in I^\leq} \left[\max \left[f_j(\mathbf{x}^2)/w_j^h - \bar{z}_j^h,\ 0 \right] \right], \\ & f_i(\mathbf{x}^1) - f_i(\mathbf{x}^h),\ (i \in I^=), \\ & f_i(\mathbf{x}^1) - \varepsilon_i^h,\ (i \in I^>), \\ & g_l(\mathbf{x}^1),\ (l = 1, \ldots, m) \Big\}. \end{aligned}$$

Notice that the weighting coefficients w_i^h, $i \in I^< \cup I^\leq$, are not included in multiobjective problem (5.12.1), but are taken into account inside the MPB method.

As explained in Section 2.2, the minimization of the improvement function takes place iteratively. For details, see also Mäkelä (1993) and Miettinen and Mäkelä (1995, 1998a).

5.12.3. Scalar Subproblem

In the scalar version of NIMBUS, after the classification, a *scalar subproblem*

$$
\begin{aligned}
&\text{minimize} && \max_{\substack{i\in I^{<}\\ j\in I^{\leq}}} \left[w_i^h(f_i(\mathbf{x})-z_i^\star), w_j^h \max\left[f_j(\mathbf{x})-\bar{z}_j^h,\ 0\right]\right] \\
&\text{subject to} && f_i(\mathbf{x}) \leq f_i(\mathbf{x}^h) \quad \text{for all} \quad i \in I^{<} \cup I^{\leq} \cup I^{=}, \\
&&& f_i(\mathbf{x}) \leq \varepsilon_i^h \quad \text{for all} \quad i \in I^{>}, \\
&&& \mathbf{x} \in S
\end{aligned} \tag{5.12.2}
$$

is formed (see also Miettinen and Mäkelä (1996b, 1998a) and Miettinen et al. (1996b)), where $z_i^\star$ for $i \in I^{<}$ are components of the ideal objective vector (assumed to be known globally).

Notice that problem (5.12.2) is nondifferentiable but has one objective function. It can be solved by any method for nondifferentiable single objective optimization, for example, by efficient bundle methods (see Mäkelä and Neittaanmäki (1992, pp. 112–137)).

Scalar subproblem (5.12.2) can be formulated in an alternative form:

$$
\begin{aligned}
&\text{minimize} && \max_{i\in I^{<}\cup I^{\leq}} \left[w_i^h \max\left[f_i(\mathbf{x}), \bar{z}_i^h\right]\right] \\
&\text{subject to} && f_i(\mathbf{x}) \leq f_i(\mathbf{x}^h) \quad \text{for all} \quad i \in I^{<} \cup I^{\leq} \cup I^{=}, \\
&&& f_i(\mathbf{x}) \leq \varepsilon_i^h \quad \text{for all} \quad i \in I^{>}, \\
&&& \mathbf{x} \in S,
\end{aligned} \tag{5.12.3}
$$

where the aspiration level is constant $\bar{z}_i^h = z_i^\star$ for $i \in I^{<}$. This formulation seems somewhat simpler but the idea is the same. Subproblems (5.12.2) and (5.12.3) do not, however, produce identical results because of scaling differences. In the following, we refer to problem (5.12.2) as the scalar subproblem but problem (5.12.3) could equally be used instead.

We prove in Subsection 5.12.5 that the solutions of the vector and the scalar subproblem are weakly Pareto optimal under certain conditions.

5.12.4. NIMBUS Algorithm

The solution of vector subproblem (5.12.1) or scalar subproblem (5.12.2) is denoted by $\hat{\mathbf{x}}^h$. If the decision maker does not like the objective vector $\mathbf{f}(\hat{\mathbf{x}}^h)$ for some reason, (s)he can explore other solutions between $\mathbf{x}^h$ and $\hat{\mathbf{x}}^h$. This means that we calculate a search direction $\mathbf{d}^h = \hat{\mathbf{x}}^h - \mathbf{x}^h$ and provide more solutions

by taking steps of different sizes in this direction. The step-size is determined by the decision maker as in the GDF method. Objective vectors $\mathbf{f}(\mathbf{x}^h + t\mathbf{d}^h)$ are calculated with different values of t ($0 \leq t \leq 1$). Their weakly Pareto optimal counterparts are presented to the decision maker, who then selects the most satisfying solution among the alternatives.

A detailed algorithm of the NIMBUS method is given below. The same algorithm is valid for both of the NIMBUS versions. Note that the decision maker must be ready to give up something in order to attain improvement for some other objective functions. The search procedure stops if the decision maker does not want to improve any objective function value.

(1) Select subproblem (5.12.1) or (5.12.2) to be used in the continuation. Choose a starting point $\mathbf{x} \in \mathbf{R}^n$ and project it onto the feasible region by solving auxiliary problem (5.12.4). Denote the new point by $\mathbf{x}^0$. Calculate its weakly Pareto optimal counterpart $\mathbf{x}^1$ by setting $I^< = \{1, \ldots, k\}$ and by solving the selected subproblem. Set the iteration counter $h = 1$.

(2) Ask the decision maker to divide the objective functions into the classes $I^<$, $I^\leq$, $I^=$, $I^>$, and I° at the point $\mathbf{z}^h = \mathbf{f}(\mathbf{x}^h)$ such that $I^= \cup I^> \cup I^\circ \neq \emptyset$ and $I^< \cup I^\leq \neq \emptyset$. If either of the unions is empty, go to step (9). Ask the decision maker for the aspiration levels $\bar{z}_i^h$ for $i \in I^\leq$ and the upper bounds ε_i^h for $i \in I^>$. Ask also for the optional weighting coefficients $w_i^h > 0$ for $i \in I^< \cup I^\leq$, summing up to one.

(3) Calculate $\hat{\mathbf{x}}^h$ by solving the subproblem. If $\hat{\mathbf{x}}^h = \mathbf{x}^h$, ask the decision maker whether (s)he wants to try another classification. If yes, set $\mathbf{x}^{h+1} = \mathbf{x}^h$, $h = h + 1$, and go to step (2); if no, go to step (9).

(4) Now $\hat{\mathbf{x}}^h$ is a new solution. Let us denote $\hat{\mathbf{z}}^h = \mathbf{f}(\hat{\mathbf{x}}^h)$. Present $\mathbf{z}^h$ and $\hat{\mathbf{z}}^h$ to the decision maker. If the decision maker wants to see different alternatives between $\mathbf{z}^h$ and $\hat{\mathbf{z}}^h$, set $\mathbf{d}^h = \hat{\mathbf{x}}^h - \mathbf{x}^h$ and go to step (6). If the decision maker prefers $\mathbf{z}^h$, set $\mathbf{x}^{h+1} = \mathbf{x}^h$ and $h = h + 1$, and go to step (2).

(5) The decision maker wants now to continue from $\hat{\mathbf{z}}^h$. If $I^< \neq \emptyset$, set $\mathbf{x}^{h+1} = \hat{\mathbf{x}}^h$, $h = h + 1$, and go to step (2). Otherwise ($I^< = \emptyset$), the weak Pareto optimality must be guaranteed by setting $I^< = \{1, \ldots, k\}$ and solving the subproblem. Let the solution be $\check{\mathbf{x}}^h$. Set $\mathbf{x}^{h+1} = \check{\mathbf{x}}^h$ and $h = h + 1$, and go to step (2).

(6) Ask the decision maker to specify the desired number of alternatives P and calculate vectors $\mathbf{f}(\mathbf{x}^h + t_j\mathbf{d}^h)$, $j = 2, \ldots, P-1$, where $t_j = \frac{j-1}{P-1}$.

(7) Produce weakly Pareto optimal objective vectors from the vectors above by solving auxiliary problem (5.12.5).

(8) Present the P alternatives to the decision maker and let her or him choose the most preferred one among them. Denote the corresponding decision vector by $\mathbf{x}^{h+1}$ and set $h = h + 1$. If the decision maker wants to continue, go to step (2).

(9) Check the Pareto optimality of $\mathbf{x}^h$ by solving auxiliary problem (2.10.2) of Part I with $\mathbf{x}^h$ as $\mathbf{x}^*$. Let the solution be $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\varepsilon}})$.

(10) Stop with the final solution $\tilde{\mathbf{x}}$.

The projection in step (1) is connected to the fact that most solvers necessitate feasible starting points. Thus, it is a more implementational than algorithmic matter. Let us mention that, for example, if the feasible region consists of inequality constraints $g_i(\mathbf{x}) \leq 0$ for $i = 1, \ldots, m$, any starting point can be projected onto the feasible region by solving the auxiliary problem

$$\begin{aligned} &\text{minimize} && \max\left[0, g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x})\right] \\ &\text{subject to} && \mathbf{x} \in \mathbf{R}^n. \end{aligned} \tag{5.12.4}$$

The justification of step (5) is given in the optimality results of Subsection 5.12.5. If we only employ the class $I^{\leq}$ to minimize functions (and the class $I^{<}$ is empty) we do not necessarily stay within the weakly Pareto optimal set. In this case we project the obtained result onto the weakly Pareto optimal set. This is acceptable according to assumption 1.

The intermediate solutions between $\mathbf{z}^h$ and $\hat{\mathbf{z}}^h$ are not necessarily weakly Pareto optimal. That is why they have to be projected onto the weakly Pareto optimal set. A practical way of doing so is to employ the results of Corollary 3.5.6 and solve the auxiliary problem

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots,k} \left[f_i(\mathbf{x}) - f_i(\mathbf{x}^h + t_j \mathbf{d}^h)\right] \\ &\text{subject to} && \mathbf{x} \in S \end{aligned} \tag{5.12.5}$$

for every $j = 2, \ldots, P-1$. This treatment works for convex as well as nonconvex problems. An alternative method can be applied if the vector subproblem is used and the problem is convex (see Miettinen and Mäkelä (1995)). In this case weak Pareto optimality can be guaranteed by solving the vector subproblem with $I^{<} = \{1, \ldots, k\}$ starting from each intermediate solution.

Since the Pareto optimality of the solutions produced cannot be guaranteed (see Subsection 5.12.5), we check the final solution in the end by solving an additional problem introduced in Theorem 2.10.3 of Part I. As the decision maker was assumed to prefer less to more, we can presume that (s)he is satisfied with the Pareto optimal final solution even where it was not her or his choice. For clarity of notation, it is not stated in the algorithm that the decision maker may check Pareto optimality at any time during the solution process. Then, problem (2.10.2) of Part I is solved with the current solution as $\mathbf{x}^*$.

Note that, if scalar subproblem (5.12.2) is employed in the algorithm, we have to calculate the components of the ideal objective vector $\mathbf{z}^*$ in the first step. However, presenting $\mathbf{z}^*$ to the decision maker gives valuable information about the problem in both NIMBUS versions.

We must remember that we cannot guarantee global optimality. If the solution obtained is not completely satisfactory, one can always solve the problem again from a different starting point. This action is also advised if the decision maker has to stop the solution process with $\hat{\mathbf{x}}^h = \mathbf{x}^h$ after step (3).

It is also possible to improve the algorithm in step (3) to avoid the case $\hat{\mathbf{x}}^h = \mathbf{x}^h$. If the upper bounds specified by the decision maker are too tight, one can use them as a reference point and project them with (5.12.5) onto the (weakly) Pareto optimal set. Showing the new solution to the decision maker provides her or him with information concerning the possibilities and the limitations of the problem, and some dead ends can be avoided as well.

Unlike some other methods based on classification, the success of the solution process does not depend entirely on how well the decision maker manages in specifying the classification and the appropriate parameter values. It is important that the classification is not irreversible. Thus, no irrevocable damage is caused in NIMBUS if the solution $\mathbf{f}(\hat{\mathbf{x}}^h)$ is not what was expected. The decision maker is free to go back or explore intermediate points. (S)he can easily get to know the problem and its possibilities by specifying, for example, loose upper bounds and examining intermediate solutions. NIMBUS is indeed learning-oriented.

5.12.5. Optimality Results

First, we state two theoretical results concerning the optimality of the solutions of vector subproblem (5.12.1) and scalar subproblem (5.12.2).

Theorem 5.12.1. The Pareto optimal solution of vector subproblem (5.12.1) is weakly Pareto optimal (to the original multiobjective optimization problem) if the set $I^<$ is nonempty.

Proof. Let us denote the feasible region of vector subproblem (5.12.1) by $\hat{S}$. Let $\mathbf{x}^* \in \hat{S}$ be a Pareto optimal solution of the vector subproblem with some sets $I^<$, $I^\leq$, $I^=$, $I^>$ and I°, where $I^< \neq \emptyset$. In other words, there does not exist another decision vector $\mathbf{x} \in \hat{S}$ such that $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)$ for all $i \in I^<$ and $\max_{j \in I^\leq} \left[\max\left[f_j(\mathbf{x}) - \bar{z}_j^h,\ 0\right]\right] \leq \max_{j \in I^\leq} \left[\max\left[f_j(\mathbf{x}^*) - \bar{z}_j^h,\ 0\right]\right]$ and at least one of the inequalities is strict.

Let us assume that $\mathbf{x}^*$ is not weakly Pareto optimal for the original problem. This means that there exists a decision vector $\mathbf{x}^\circ \in S$ such that $f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$.

Because $\mathbf{x}^*$ is a feasible solution of problem (5.12.1), we have $f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*) \leq f_i(\mathbf{x}^h)$ for $i \in I^=$ and $f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*) \leq \varepsilon_i^h$ for $i \in I^>$. Thus, also $\mathbf{x}^\circ \in \hat{S}$.

For all $i \in I^\leq$ is valid $f_i(\mathbf{x}^\circ) - \bar{z}_i^h < f_i(\mathbf{x}^*) - \bar{z}_i^h$. It implies that $\max\left[f_i(\mathbf{x}^\circ) - \bar{z}_i^h,\ 0\right] \leq \max\left[f_i(\mathbf{x}^*) - \bar{z}_i^h,\ 0\right]$ for all $i \in I^\leq$, and, further,

$$\max_{i \in I^\leq} \Big[\max\left[f_i(\mathbf{x}^\circ) - \bar{z}_i^h,\ 0\right]\Big] \leq \max_{i \in I^\leq} \Big[\max\left[f_i(\mathbf{x}^*) - \bar{z}_i^h,\ 0\right]\Big].$$

While, in addition,

$$f_i(\mathbf{x}^\circ) < f_i(\mathbf{x}^*),$$

for all $i \in I^< \neq \emptyset$, the point $\mathbf{x}^*$ cannot be a Pareto optimal solution of the vector subproblem. This contradiction implies that $\mathbf{x}^*$ must be weakly Pareto optimal. The proof is also valid if some of the classes $I^\leq$, $I^=$, $I^>$ or $I^\diamond$ are empty as long as $I^> \cup I^\diamond \neq \emptyset$. □

Theorem 5.12.2. The solution of scalar subproblem (5.12.2) is weakly Pareto optimal if the set $I^<$ is nonempty.

Proof. Let us denote the objective function of scalar subproblem (5.12.2) by $f(\mathbf{x})$ to be minimized and the feasible region by $\hat{S}$. Let $\mathbf{x}^* \in \hat{S}$ be a solution of the scalar subproblem with some sets $I^<$, $I^\leq$, $I^=$, $I^>$, and $I^\diamond$, where $I^< \neq \emptyset$. In other words, $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \hat{S}$.

Let us assume that $\mathbf{x}^*$ is not weakly Pareto optimal. This means that there exists a vector $\mathbf{x}^\diamond \in S$ such that $f_i(\mathbf{x}^\diamond) < f_i(\mathbf{x}^*)$ for all $i = 1, \ldots, k$.

Note in the following that all the weighting coefficients are strictly positive.

Because $\mathbf{x}^* \in \hat{S}$, we have $f_i(\mathbf{x}^\diamond) < f_i(\mathbf{x}^*) \leq f_i(\mathbf{x}^h)$ for $i \in I^< \cup I^\leq \cup I^=$ and $f_i(\mathbf{x}^\diamond) < f_i(\mathbf{x}^*) \leq \varepsilon_i$ for $i \in I^>$. Thus, also $\mathbf{x}^\diamond \in \hat{S}$.

Since $z_i^\star \leq f_i(\mathbf{x}^\diamond) < f_i(\mathbf{x}^*)$ for all $i \in I^< \neq \emptyset$, we have $f_i(\mathbf{x}^*) - z_i^\star > 0$ for all $i \in I^<$. With $w_i > 0$ we also have $w_i(f_i(\mathbf{x}^*) - z_i^\star) > 0$ for all $i \in I^<$.

Let us consider

$$f(\mathbf{x}^\diamond) = \max_{\substack{i \in I^< \\ j \in I^\leq}} \left[w_i(f_i(\mathbf{x}^\diamond) - z_i^\star), w_j \max\left[f_j(\mathbf{x}^\diamond) - \bar{z}_j^h,\ 0\right] \right].$$

The maximum can be attained either in the class $I^<$ or in $I^\leq$ (or, naturally, in both of them). In the first case we have

$$f(\mathbf{x}^\diamond) = w_i(f_i(\mathbf{x}^\diamond) - z_i^\star) < w_i(f_i(\mathbf{x}^*) - z_i^\star) \leq f(\mathbf{x}^*)$$

for some $i \in I^<$. The latter case has two different alternatives. Firstly,

$$f(\mathbf{x}^\diamond) = w_j \max\left[f_j(\mathbf{x}^\diamond) - \bar{z}_j^h,\ 0\right] = 0 < w_i(f_i(\mathbf{x}^*) - z_i^\star) \leq f(\mathbf{x}^*)$$

for some $j \in I^\leq$ and for all $i \in I^<$. Secondly,

$$f(\mathbf{x}^\diamond) = w_j \max\left[f_j(\mathbf{x}^\diamond) - \bar{z}_j^h,\ 0\right] = w_j(f_j(\mathbf{x}^\diamond) - \bar{z}_j^h) < w_j(f_j(\mathbf{x}^*) - \bar{z}_j^h) \leq f(\mathbf{x}^*)$$

for some $j \in I^\leq$.

In conclusion, we can state that the point $\mathbf{x}^*$ cannot be a solution of the scalar subproblem. This contradiction implies that $\mathbf{x}^*$ must be weakly Pareto optimal. The proof is also valid if some of the classes $I^\leq$, $I^=$, $I^>$ or $I^\diamond$ are empty as long as $I^> \cup I^\diamond \neq \emptyset$. □

The following optimality result is common for the scalar and the vector subproblem (even the proofs can be combined).

Theorem 5.12.3. Any Pareto optimal solution can be found with an appropriate classification in problems (5.12.1) and (5.12.2).

Proof. Let $\mathbf{x}^* \in S$ be Pareto optimal. Let us assume that there does not exist a classification such that $\mathbf{x}^*$ is a solution of the vector or the scalar subproblem. Let us suppose that we have the current NIMBUS solution $\mathbf{x}^h$ and the corresponding $\mathbf{z}^h$ available.

Let us choose $\mathbf{f}(\mathbf{x}^*)$ as a reference point. This means that we choose $\bar{z}_i = f_i(\mathbf{x}^*)$ for those indices where $z_i^h > f_i(\mathbf{x}^*)$ and set $i \in I^{\leq}$. Further, we set $\varepsilon_i = f_i(\mathbf{x}^*)$ for indices $i \in I^{>}$ satisfying $z_i^h < f_i(\mathbf{x}^*)$. Finally, the set $I^{=}$ consists of indices where $z_i^h = f_i(\mathbf{x}^*)$. This setting is possible because $\mathbf{x}^*$ is assumed to be Pareto optimal and $\mathbf{x}^h$ is weakly Pareto optimal according to Theorems 5.12.1 and 5.12.2 and the structure of the NIMBUS algorithm. That is why $I^{\leq} \neq \emptyset$ and $I^{=} \cup I^{>} \neq \emptyset$. In addition, we set $w_i = 1$ for $i \in I^{\leq}$.

Because $\mathbf{x}^*$ is not a solution of the vector or the scalar subproblem, there exists another point $\mathbf{x}^\circ \in S$ that is a solution, meaning that

$$\max_{j \in I^{\leq}} \left[\max \left[f_j(\mathbf{x}^\circ) - f_j(\mathbf{x}^*),\ 0 \right] \right] < \max_{j \in I^{\leq}} \left[\max \left[f_j(\mathbf{x}^*) - f_j(\mathbf{x}^*),\ 0 \right] \right] = 0.$$

Thus, $\max [f_j(\mathbf{x}^\circ) - f_j(\mathbf{x}^*),\ 0] < 0$ for every $j \in I^{\leq}$. In other words, we have $f_j(\mathbf{x}^\circ) < f_j(\mathbf{x}^*)$ for every $j \in I^{\leq}$. Because $\mathbf{x}^\circ$ is a solution of problems 5.12.1 and 5.12.2, it must be feasible. In other words, we have $f_i(\mathbf{x}^\circ) \leq f_i(\mathbf{x}^*)$ for $i \in I^{=} \cup I^{>}$. Here we have a contradiction to the assumption that $\mathbf{x}^*$ is Pareto optimal. This completes the proof and $\mathbf{x}^*$ must be a solution of the vector and the scalar subproblem. □

The MPB method is an essential element of the vector version of NIMBUS. It can also be used for solving the scalar version of NIMBUS. The performances of the two versions are comparable if the same solver is employed in both of them. Thus, the optimality of the solutions produced by the MPB method is an important fact to consider. The optimality of the solutions of the MPB method was handled in Subsection 2.2.3. Accordingly, in theory only the substationarity of the solutions of the MPB method is guaranteed for general multiobjective optimization problems. For fluency, we have thus far referred and shall continue referring to the solutions as weakly Pareto optimal. Note, on the other hand, that a global single objective optimizer can be employed with the scalar version to produce globally (weakly) Pareto optimal solutions.

5.12.6. Comparison of the Two Versions

The vector and the scalar versions of NIMBUS differ in the form of the subproblem used. The origin of the development of the scalar version lies in the drawbacks discovered in the vector version.

Theoretically, the solution of the vector subproblem has to be Pareto optimal in order to guarantee weakly Pareto optimal solutions to the original multiobjective optimization problem. This assumption is quite demanding. With the scalar subproblem we do not have problems of this kind. Further, the vector version needs a special solution tool – MPB. In addition to this limitation, the

role of the weighting coefficients is not commensurable between the classes $I^<$ and $I^\leq$. This implies that the controllability of the method suffers.

The advantage of having a single objective function in the scalar version is that we can employ any efficient optimization routine of nondifferentiable optimization. This gives more generality and applicability to the method. Furthermore, in the scalar subproblem, we treat the functions in $I^<$ and $I^\leq$ in a consistent way and, thus, the roles of the weighting coefficients are also identical. In all, this means that the decision maker can better direct the solution process.

Notice that in addition to the difference in the objective functions of the subproblems, there is also deviation in the constraint part. Due to the goal of the classes $I^<$ and $I^\leq$, we have to make sure that the values of these functions do not increase. This is the reason for modifying the constraints of scalar subproblem (5.12.2).

In the vector subproblem, the MPB method does not allow increment in $I^<$. However, there is no guarantee that the values of the functions in $I^\leq$ could not increase. It is clear that including additional constraints in an optimization problem increases its computational complexity. Because the increasing feature occurs very rarely in the vector version, no additional constraints have been used in order to emphasize computational efficiency. Thus, either we pay the price of additional computational costs or take the risk of increment (depending on the classification).

On the one hand, the calculation of the ideal objective vector used in the scalar version also needs computational effort. On the other hand, the ideal objective vector can provide supporting information for the decision maker in any kind of multiobjective solution process.

A numerical comparison of the two versions of NIMBUS is reported in Miettinen and Mäkelä (1996b, 1998a) with versatile multiobjective optimization problems. The standards of comparison chosen are computational efficiency and the opinion of the decision maker concerning the controllability of the different versions. The efficiency can be measured by the number of times the subroutine containing the objective functions is called. The controllability side must be elicited from the decision maker. It is measured in the form of a rating (between 1 and 5).

The numerical tests indicate that the scalar version obeys the decision maker better, whereas the vector version is computationally more efficient. However, it is important to note that the classifications employed affect considerably the performance of the NIMBUS versions. In any case, the user has to choose between controllability and computational efficiency when selecting a solution method.

5.12.7. Comments

The NIMBUS method has not been developed to converge in the traditional sense. While the method does not assume the existence of any underlying value function, no explicit convergence results can be put forward on the basis of the assumptions about the properties of the function. In particular, the intention has been to release the decision maker from the assumption of an underlying value function. What is important is that the method satisfies two desirable properties of interactive methods: not to place too demanding assumptions on the decision maker or the information exchanged, and to be able to find (weakly) Pareto optimal solutions quickly and efficiently.

The aim has been to formulate a method where the decision maker can easily explore the (weakly) Pareto optimal set. When the decision maker no longer wants to change any objective function value and the solution process stops, the solution is then optimal.

An important factor is that the final solution is always Pareto optimal because of the structure of the algorithm. In addition, all the intermediate points are at least substationary points and they can be projected onto the Pareto optimal set, if so desired.

The method is ad hoc in nature, since the existence of a value function would not directly advise the decision maker how to act to attain her or his desires. A value function could only be used to compare different alternatives.

The possibility of interchanging the roles of the objective and the constraint functions has been mentioned thus far in connection with some methods. This is easy to carry out also in NIMBUS because the class $I^>$ is nothing but constraints with upper bounds. One can even go that far as to formulate all the constraint functions as objective functions and modify their upper bounds or roles during the solution process from one iteration to the other.

5.12.8. Implementations

The NIMBUS algorithm was originally implemented in the mainframe environment at the University of Jyväskylä, Finland. This approach is suitable for even large-scale problems, but the lack of a flexible user interface decreases its usability. It is evident that the user interface plays a crucial role in realizing interactive algorithms.

An alternative is to use microcomputers to develop a functional user interface by paying the price of reduced computational capacity. However, both the mainframe and the microcomputer environment share weaknesses in common from the viewpoint of both the user and the developer of the implementation. As far as the user is concerned, the system and some specific compilers have to be installed. It is limiting that the programs are appropriate only for certain computer environments and operating systems. For the developer the delivery of the software updates is laborious. Implementing and keeping up separate versions for different environments requires also extra effort.

The strengths of the mainframe and the microcomputer environment are the computational efficiency and graphical user interface, respectively. On the one hand, the readiness to combine them and, on the other hand, the rapid expansion in the use of the Internet and the World-Wide Web (WWW) have motivated the development of a WWW-NIMBUS system in the Internet (see Miettinen and Mäkelä (1998b)).

Via the Internet we can centralize the computing to one server computer (at the University of Jyväskylä) and the WWW is a way of distributing the interface to the computers of each individual user. The WWW enables the realization of a graphical user interface. Thus, it is possible to implement an interactive optimization algorithm by the means of the WWW. The idea is not to demand high computing capacity or special compilers from the computer of the user. Instead, this is left to one efficient server computer. The same computer takes care of the visualization and the problem data management. This centralizing offers benefits to both the user and the software producer. The user always has the latest version of the NIMBUS method available, and the producer has only to update and develop one version. In addition, the user saves the trouble of installing the software.

The most important aspect of WWW-NIMBUS is that it is easily accessible and available to any Internet user (http://nimbus.math.jyu.fi/). No special tools, compilers or software besides a WWW browser are needed. The system is independent of the computer and the operating system used.

The WWW environment enables the possibility of graphical classification and graphical visualization of the alternatives. They both support the decision maker in getting to know the problem and finding better solutions.

5.12.9. Applications

In Miettinen (1994), two academic problems and a state-constrained optimal control problem concerning an elastic string are solved by the vector version of NIMBUS. The vector version is applied to solve an academic nondifferentiable test problem and a river pollution problem in Miettinen and Mäkelä (1995). A nondifferentiable version of the pollution problem is dealt with in Miettinen and Mäkelä (1997). A structural design problem is solved by the vector version of NIMBUS in Miettinen et al. (1996a). In Miettinen and Mäkelä (1996b, 1998a), a water quality management problem is solved by both the vector and the scalar versions.

An optimal control problem related to the continuous casting of steel is solved by the vector version in Miettinen (1994) and by the scalar version in Miettinen et al. (1996b). This problem is an example of the case where the modelling phase ends up with an empty feasible region. The so-called technological constraints are so tight that there does not exist any feasible solution. When this happens, the constraints can be treated as objective functions with the original objective function(s) thus forming a multiobjective optimization

problem. One of the goals is then to find a solution as close to the feasible region as possible.

5.12.10. Concluding Remarks

NIMBUS is one of the few efficient, interactive methods especially developed for solving nondifferentiable multiobjective optimization problems. Naturally, differentiable problems can be solved as well. In two different versions of NIMBUS, the decision maker moves around the weakly Pareto optimal set and expresses iteratively her or his desires by specifying those objectives whose values should improve and those whose values are allowed to deteriorate with the help of five available classes. The selection of the most preferred alternative from a given set is also possible. The questions posed to the decision maker are not demanding. The method aims at being flexible and the decision maker can select to what extent (s)he exploits the versatile possibilities of the method. The calculations are not too massive, either. The use of efficient bundle methods as the underlying nondifferentiable optimizers is recommended (see Mäkelä and Neittaanmäki (1992, pp. 138–143)).

In NIMBUS, the decision maker is free to explore the (weakly) Pareto optimal set and also to change her or his mind if necessary. Previous acts do not limit the movements. The decision maker can also extract undesirable solutions from further consideration. Naturally, the decision maker does not have to employ all of the five classes if (s)he feels uncomfortable with some of them. However, it is important to provide the decision maker with alternative courses of action.

The classification of the functions and the specification of the appropriate parameter information does not necessarily have to succeed as well as in some other classification-based methods. The reason is that intermediate solutions can be examined and, thus, more information about the problem obtained. This makes the method more flexible. In addition, the decision maker can cancel any classification step because nothing is irreversible. A further advantage is that not all the objective functions have to be classified. This is not possible in other classification-based methods. In NIMBUS, the decision maker can set some objective function free and examine what happens to the other objectives.

Even though the Pareto optimality of the solutions produced cannot be guaranteed, at least the final solution is Pareto optimal. If it is important to the user that the intermediate solutions are Pareto optimal, they can be projected. However, this adds to the computational costs. The fact that the Pareto optimality of the solutions produced is not guaranteed automatically is at least partly compensated for by the computational efficiency of the underlying optimizer.

From numerical experiments we can conclude that of the two NIMBUS versions available the scalar version is more controllable and dirigible to the decision maker. The advantage of the vector version is its computational effi-

ciency. Thus, the user has to choose between these aspects when selecting a solution method.

Eventually, it is up to the user interface to make the most of the possibilities of the method and provide them to the user. When the first WWW version of the NIMBUS algorithm was implemented in 1995 it was a pioneering interactive optimization system in the Internet. The realization is based on the ideas of centralized computing and a distributed interface.

Naturally, there are many challenges in the further development of the NIMBUS method and its implementations. One of the challenges, applicable to software development in general, is to create illustrative and easy-to-use user interfaces. If the interface is able to adapt to the decision maker's style of making decisions and is of help in analyzing the alternatives and results, and can perhaps give suggestions or advice, then the interface may even overcome some of the deficiencies of the method itself.

5.13. Other Interactive Methods

The number of existing interactive methods is large. That is why it is neither the purpose nor practical nor possible to discuss all of them here. Nevertheless, in addition to those presented in the previous sections, some methods are listed below. Only the basic concepts and ideas of the methods are mentioned together with references. The methods are roughly divided according to their basis in goal programming, in weighted metrics, in reference points, and in miscellaneous ideas. Some methods for linear problems are included because of their interesting basic ideas or because they are referred to in connection with method comparisons in Section 1.2 of Part III.

5.13.1. Methods Based on Goal Programming

A rather straightforward extension of goal programming into an interactive form is presented in Masud and Hwang (1981). The method is called the interactive sequential goal programming (ISGP). The interactive multiple goal programming (IMGP) method, described in Nijkamp and Spronk (1980) and Spronk (1990), has also been created to combine the flexibility of goal programming and the robustness of interactive approaches. The decision maker indicates which objective value(s) should be improved and either revises the aspiration levels of the corresponding goals or the problem is automatically modified with additional constraints.

The sequential multiobjective problem solving (SEMOPS) technique is briefly outlined in Monarchi et al. (1973). Five types of goal specifications in the form of points and intervals are allowed. A different measure of deviation is utilized for each type. (For example, if the goal is of the form $f_i(\mathbf{x}) \leq \bar{z}_i$, then the corresponding measure of deviation is $\delta_i = f_i(\mathbf{x})/\bar{z}_i$.) At each iteration, a

subset of deviations is summed up and then minimized. The decision maker may change that subset and specify new aspiration levels. Unfortunately, the solutions are not guaranteed to be Pareto optimal. A related method, called the sequential information generator for multiple objective problems (SIGMOP), is introduced in Monarchi et al. (1976). SIGMOP is a flexible method where the decision maker can alter aspiration levels and weighting coefficients as (s)he separates attainable solutions from among the desired ones. As an application, a pollution problem in water resources is solved by the SEMOPS and the SIGMOP methods in the references mentioned.

The ideas of goal programming, the ε-constraint method and trade-offs are combined in the direction-searching method proposed in Masud and Zheng (1989). The method aims at reducing the cognitive burden on the decision maker while not increasing computational complexity. The algorithm is illustrated by a numerical example. The properties of the method are also compared with those of several other interactive methods.

The general purpose interactive goal programming algorithm is suggested in Tamiz and Jones (1997b). An interactive goal programming algorithm for nonlinear problems based on different norms and updating the aspiration levels is presented in Weistroffer (1983).

5.13.2. Methods Based on Weighted Metrics

The idea of the method in Moldavskiy (1981) is to form a grid in the space of the weighting vectors and to map this grid onto the Pareto optimal set. Weighted L_p-metrics are used as scalarizing functions to produce a representation of the Pareto optimal set. The decision maker can contract the space of the weighting vectors until the most satisfactory solution is obtained.

A method based on sensitivity analysis and the weighted Tchebycheff metric is presents in Diaz (1987), where the effects of changing aspiration levels are studied by sensitivity analysis. The method in Sunaga et al. (1988) utilizes also the weighted Tchebycheff metric. It transforms the constrained min-max problem into a series of (differentiable) unconstrained problems by penalty functions.

The interactive cutting-plane algorithm is presented in Loganathan and Sherali (1987) with applications. The idea is to maximize the underlying value function. The weighted Tchebycheff metric is utilized with marginal rates of substitution as weighting coefficients. The convergence of the algorithm is also treated.

The method proposed in M'silti and Tolla (1993) combines features from the ε-constraint method and the augmented weighted Tchebycheff metric. The global Pareto optimality of the solutions obtained is checked.

The method of the displaced ideal for MOLP problems, described in Zeleny (1973, 1974, 1976), can be characterized as an interactive extension of the method of weighted metrics. A subset of the Pareto optimal set is obtained by minimizing the distance between the ideal objective vector and the feasible

objective region by the weighted L_p-metrics with altered exponents p. The subset is reduced by moving the reference point towards the feasible objective region until the subset of the Pareto optimal solutions is small enough for the decision maker to select the most preferred solution. The method is based on empirical studies of the decision maker's behaviour.

The distance to the utopian objective vector is minimized by weighted L_p-metrics in Köksalan and Moskowitz (1994). This interactive method is based on determining the weighting coefficients according to the preferences of the decision maker. The preference information is obtained from pairwise comparisons.

Ways of approaching discrete multiattribute decision analysis problems have been included in the method introduced in Kok and Lootsma (1985). The ideal objective vector is used as a reference point. Pairwise comparison methods are applied between the reference point and the (possibly approximated) nadir objective vector. The distances are measured by solving the augmented weighted Tchebycheff problem.

An interesting method is suggested in Kaliszewski et al. (1997). In this hybrid interactive decision making technique, the decision maker can select what kind of information to specify. (S)he can either classify the objective functions or specify upper bounds on global trade-offs. New properly Pareto optimal solutions are generated by modified weighted Tchebycheff problem (3.4.9) according to the results derived in Subsection 3.4.6. This is the way of taking the bounds on the global trade-offs into account.

5.13.3. Methods Based on Reference Points

Multiple reference points and a gradient projected method are bases of the method of Costa and Clímaco (1994) for MOLP problems. The method is related to Pareto race (in Subsection 5.10.3) but it utilizes parallel processing when handling several reference points simultaneously.

The method of Wierzbicki forms the basis of the interactive reference point methods introduced in Bogetoft et al. (1988). The multiobjective optimization problem is assumed to be convex. Karush-Kuhn-Tucker multiplier information is presented to the decision maker to guide the specification of new reference points. Several different modifications are also presented and their convergence properties are studied.

In the method presented in Tapia and Murtagh (1989), the decision maker is asked to express preferential desires to attain her or his reference point. So-called preference criteria are formed from this information. These preference criteria are then used as a reference point in the achievement function. The authors also report some encouraging numerical experiments.

A method combining the ideas of reference points and measuring distances is suggested in Hallefjord and Jörnsten (1986). After the decision maker has specified the reference point, the distance between it and the feasible objective

region is minimized by an entropy function. The mathematical background of the method is widely handled in the reference.

The method presented in Weistroffer (1982) assumes that the decision maker specifies required values or maximum-achievement levels. The surplus is then maximized to the Pareto optimal set. Further, the methods in Narula and Weistroffer (1989b) and Weistroffer (1984, 1987) expect the decision maker to provide both the required and desired values for every objective function. Then an achievement function is optimized. The required and the desired values are modified until the most preferred solution is obtained. Some convergence results are also dealt with.

The bi-reference procedure presented in Michalowski and Szapiro (1992) has been developed for MOLP problems. The decision maker is asked to specify the worst acceptable objective vector, and a search direction is obtained as the difference between the worst and the ideal objective vectors. As long a step as possible is taken in that direction and the decision maker is asked to divide the objective functions into three classes (to be improved, to be kept unchanged and to be relaxed). Then the worst and the ideal objective vectors are replaced and the procedure continues until no significant improvement is achieved. The performance of the bi-reference procedure is compared with other interactive procedures by solving some test examples from the literature. At least in those examples the procedure manages quite well.

An extreme point method for MOLP problems is described in Kirilov and Vassilev (1997). It is based on the reduction of the weighting space and is related to the Tchebycheff method. The decision maker can compare different solution alternatives and guide the solution process by specifying reference points.

5.13.4. Methods Based on Miscellaneous Ideas

An interactive extension of the weighting method is presented in Steuer (1986, pp. 394–399). Many of its ideas are related to those of the Tchebycheff method. The set of the weighting vectors is reduced according to the choices of the decision maker. Weighting vectors are generated randomly from the reduced space and filtered to obtain a well dispersed set. Below, this approach will be called the method of Steuer.

Another interactive method based on the weighting method is introduced in Hussein and El-Ghaffar (1996). It can handle convex problems and is based on solving systems of equations formed according to the Karush-Kuhn-Tucker type optimality conditions.

Two different interactive relaxation methods are put forward in Nakayama et al. (1980) and Lazimy (1986b). The latter is applicable to both continuous and integer problems. The methods are based on the maximization of an underlying value function in a new but equivalent form with additional constraints. Marginal rates of substitution and other estimates of the value function are required from the decision maker. Similar ingredients are utilized in

the decomposition method presented in Lazimy (1986a). It is based on the duality theory for nonlinear programming. The original problem is decomposed into a series of linear subproblems and two-attribute problems. In addition, a relaxation-projection technique, especially for bi-objective problems, is proposed in Ferreira and Geromel (1990). An application to scheduling is also handled.

An interactive algorithm with several alternative subproblems is proposed in Mukai (1980). The subproblems generate feasible directions in which the values of all the objective functions improve. The decision maker can then indicate what objective functions to improve at the expense of others, and a new direction is generated. Tools for extending the applicability of Mukai's algorithms to nondifferentiable objective functions are presented in Kiwiel (1984, 1985a, b) and Wang (1989). The ideas were applied in the MPB method in Section 2.2.

The method introduced in Roy and Mackin (1991) is based on a sequence of pairwise questions and it tries to approximate the parameters of a proxy value function.

An example of including ideas from other research areas in interactive multiobjective optimization is presented in Tapia and Murtagh (1992). The preferences of a decision maker are analyzed in MOLP problems with Markovian processes. The cardinal priority ranking of Pareto optimal solutions is a part of the method proposed.

Methods for discrete and continuous problems are combined in Slowinski (1991). A finite set of Pareto optimal points is generated and then ordinal regression is applied. The method is intended to be practical in situations where the decision maker wants to focus on a subset of Pareto optimal points at early stages of the process. Similar ideas are utilized in Bard (1986). A set of Pareto optimal solutions is generated by the ε-constraint method and ranked by means of multiattribute decision analysis. The method is demonstrated by an example concerning the selection of automation options for an upcoming space station.

The interactive step trade-off method combining ideas from the SWT method and STEM is presented in Yang et al. (1990). It utilizes trade-off rates and the division of objectives into those to be improved, those that should maintain their values and those to be impaired. In Yang and Sen (1996), the interactive step trade-off method is extended to a two-phase algorithm to include the estimation of piecewise linear value functions based on pairwise comparisons.

Possibilities of multiobjective optimization in structural mechanics are presented in Eschenauer et al. (1989). Two interactive methods are briefly described and applied to the optimization of a conical shell.

The Zionts-Wallenius (ZW) method for MOLP problems (see Wallenius and Zionts (1977) and Zionts and Wallenius (1976, 1983)) based on the weighting method and optimizing the underlying (implicitly known) value function can be generalized to convex objective functions by considering piecewise linearizations. The ZW method is extended for concave (maximization case) objective

and value functions and convex feasible regions in Roy and Wallenius (1992). A more general case of nonlinear objective functions, nonconvex feasible regions and concave value functions is also discussed. This approach uses the generalized reduced gradient method instead of the original simplex.

The method in Kim and Gal (1993) is intended for MOLP problems. It is based on the concept of a maximally changeable dominance cone and marginal rates of substitution. The effectiveness of the method is illustrated by a numerical example.

Ideas for reducing the burden on the decision maker in interactive methods are introduced in Korhonen et al. (1984) and further developed in Ramesh et al. (1988). An underlying quasiconcave value function is assumed to exist. Convex cones are formed according to the preference relations of the decision maker. The cones are formed so that the solutions in them can be dropped from further consideration, because they are dominated by some other solutions. Thus, fewer questions have to be put to the decision maker in charting the preferences. These ideas concerning convex cones can be applied equally to multiobjective optimization as to multiattribute decision analysis. The ideas are utilized, for example, in Ramesh et al. (1989a, b).

A method for complex problems with high dimensionality is proposed in Baba et al. (1988). The method uses a random optimization method and is also applicable to nondifferentiable objective functions.

The parameter space investigation (PSI) method is described briefly in Lieberman (1991b) and in more detail in Statnikov and Matusov (1996) and Steuer and Sun (1995). It has been developed for complicated nonlinear problems involving possible differential equations. Such problems occur, for example, in engineering. The method is very simple and intended to be applicable to problems where more sophisticated methods are useless. The PSI method is a naïve sampling technique rather than an optimization method. Both the constraint functions and variables are assumed to have upper and lower bounds. Thus, the feasible region is a parallelepiped. The Pareto optimal set is approximated by generating randomly uniformly distributed points between the variable bounds. Infeasible solutions are dropped as well as solutions not satisfying the upper bounds specified by the decision maker. Pareto optimal solutions are selected from this set. The sample size can be altered and the decision maker can adjust the upper bounds. The method does not assume differentiability. It works for nonconvex problems since its structure enables global search. The method contains a random number generator of its own, but it is claimed in Steuer and Sun (1995) that any generator can equally well be used. Convergence properties and the accuracy of the approximation of the Pareto optimal set assuming general Lipschitz conditions are handled in Sobol' and Levitan (1997) and Statnikov and Matusov (1996). The PSI method has its origins in the former Soviet Union, which is why most of the information about it has been published in Russian. It is said to have been applied in many fields of the national economy in Russia. One engineering application is described in Sobol (1992).

1. COMPARING METHODS

As has been stressed many times thus far, a large variety of methods exists for multiobjective optimization problems and none of them can be claimed to be superior to the others in every aspect. Selecting a multiobjective optimization method is a problem with multiple objectives itself. Thus some matters of comparison and selection between the methods are worth considering.

The theoretical properties of the methods can rather easily be compared. We summarize some of the features of the interactive methods treated in this book in a comparative table at the beginning of this chapter. However, in addition to theoretical properties, practical applicability also plays an important role in the selection of an appropriate method for the problem to be solved. The difficulty is that practical applicability is hard to determine without experience and experimentation.

More fruitful information relating to the question of method selection would likely emerge if computational applications were more extensively reported. Unfortunately, not too many actual computational applications of multiobjective optimization techniques have been published. Instead, methods have mainly been presented without computational experiences or with simple academic test problems. As it is aptly remarked in Bischoff (1986), most of the applications presented are merely proposals for applications or they deal with highly idealized problems. For most interactive methods a natural reason is the difficulty (in finding and) in testing with real decision makers. A complicating fact is also the enormous diversity of decision makers.

One more thing to keep in mind is that for the most part only successful applications are published. This means that we cannot draw a complete picture of the applicability of a method on the basis of the experiences reported.

The evident lack of benchmark-type test problems for nonlinear multiobjective optimization complicates the comparison of different methods. Naturally, some methods are useful for some problems and other methods for other types of problems. However, benchmark problems could be used to point out such behaviour.

In this section we outline some comparisons of methods reported in the literature. We also consider selected issues in deciding upon a method, including a decision tree.

K. Miettinen, *Nonlinear Multiobjective Optimization*

1.1. Comparative Table of Interactive Methods Presented

In Figure 1.1.1, we present a comparative table of the twelve interactive multiobjective optimization methods described in Chapter 5 of Part II. This can be regarded as a brief summary of these methods. However, a sceptical attitude should always be taken towards such attempts to compress matters to an extreme. The table is subjective and there is no point in even trying to deny it.

Different problems arise when one tries to put together a table of this kind. Among them are, for example, deciding what property is important enough to be included, how it should be formulated, and whether it is a positive or a negative one.

Figure 1.1.1 presents some of the properties described when the methods were introduced. They are related to the general features of the methods and their solutions. Properties concerning the nature of the assumptions set in relation to the problem to be solved have not for the most part been included. They will be handled in a decision tree in Figure 1.3.1.

The table is by no means self-contained. However, we do not explain the table here in detail but refer to the corresponding sections were the methods were presented.

The properties of the methods have been grouped into four categories according to different aspects. The first group is related to the general characteristics of the methods. The style of interaction is treated in the second group. The third group describes what kind of information is solicited from the decision maker. Finally we indicate those methods whose implementations have been mentioned in connection with the descriptions of the methods.

For clarity, explanatory comments on some of the properties listed are in order. As far as the property 'final solution Pareto optimal' is concerned, parentheses indicate methods where Pareto optimality depends on the scalarizing function used.

The distinction between 'classification of objective functions used' and 'reference points used' must be emphasized. As stated earlier, the classification and the reference points are parallel. However, classification is somehow more demanding because it necessitates pointing out both the objective functions to be relaxed and those to be improved when compared to the current solution. In Figure 1.1.1, specifying a reference point is understood to be independent of the current solution in this sense.

The implementation of the reference direction approach is in parentheses because the program can solve linear problems only.

Some of the properties have been subjectively classified into positive or negative ones. The other properties are mainly matters of taste.

	ISWT method	GDF method	SPOT	Tchebycheff method	STEM	Reference point method	GUESS method	STOM	Light beam search	Reference direction approach	RD method	NIMBUS method
+ final solution Pareto optimal	×		×	×		(×)		(×)	×	(×)	(×)	×
+ final solution weakly Pareto optimal	×		×	×	×	×	×	×	×	×	×	×
+ suitable also for nondifferentiable problems				×	×	×	×				×	×
ad hoc nature					×	×	×	×	×	×	×	×
objective functions assumed to be bounded	×	×	×	×	×		×		×	×		×
- sensitive needing consistent answers	×	×	×	×							(×)	
- computationally expensive				×					×			
- difficult questions posed	×	×	×									
+ trade-off rates provided	×		×					×				
comparison of alternatives used	×	×		×		×			×	×		×
classification of objective functions used					×			×			×	×
reference points used						×	×		×	×		
marginal rates of substitution used		×	×									
thresholds used									×			
implementation mentioned						×		×		(×)		×

Figure 1.1.1. Properties of interactive methods.

1.2. Comparisons Available in the Literature

Here we briefly mention some of the comparisons available together with a few results and some conclusions. For more detailed information, see the references cited.

1.2.1. Introduction

The comparisons have been carried out with respect to a variety of criteria. Among them are ease of use and confidence in both the solution obtained and the method used from the viewpoint of the decision maker. The rapidity of convergence and CPU time are among the criteria from the mathematical point of view. The number of Pareto optimal solutions needed to solve a problem could also serve as a comparison criterion, as pointed out in Ferreira and Machado (1996). However, such a measure of effectiveness has not generally been reported in the comparative evaluations available.

Some caution is in order when trying to judge something from the comparisons. The comparisons have been performed according to different criteria and under varied circumstances. Thus they are not fully proportional. Which method is the most suitable for a certain problem depends highly on the personality of the decision maker and on the problem to be solved.

Practical experience is especially important in evaluating the techniques with respect to criteria related to the decision maker. It is important to compare a method under a variety of circumstances so that the conclusions can be generalized. As emphasized, for example, in Hobbs et al. (1992), the appropriateness, ease of use and validity of a method must be tested with real decision makers.

Using a human decision maker does not, however, mean that the practical applicability of the method has been fully investigated. Unfortunately, few experiments have been reported with problem-related decision makers. Most of the comparative evaluations with human decision makers have involved students as the decision makers. This is understandable for practical reasons. However, this kind of a setting can be called into question. The results might have been different with real decision makers who are actually responsible for the solution obtained. Another aspect is the wide range of different problem areas and their different decision makers. Obvious examples are business-related problems that typically involve less than ten objective functions and engineering design problems with hundreds of objective functions. Further aspects to be kept in mind when testing several methods with human decision makers are the effects of learning and anchoring. Learning is related both to getting to know the problem better and to the order of the methods used whereas anchoring is related to the selection of starting points.

Instead of a human decision maker one can sometimes employ value functions in the comparisons. Value functions may be useful in evaluating theoretical performance, but such tests do not fully reflect the real usefulness of the methods. One can try to compensate for the lack of a real decision maker by employing several different value functions. If, for example, marginal rates of substitution are desired, the inconsistency and inaccurate responses of a decision maker can be imitated by multiplying them with different random numbers. These means are employed in Shin and Ravindran (1992). On the other hand, value functions cannot really help in testing ad hoc methods.

One crucial factor that can affect the performance of the methods in the comparisons is the user interface. Nothing is usually mentioned concerning the realization of the user interface in the comparisons reported. It is important to remember that one can spoil a 'good' method with a poor user interface or support a 'poor' method with a good interface. In addition to the illustration of the (intermediate) results, a good user interface also means a clear and intelligible input phase.

It is interesting to observe that most of the multiobjective optimization problems solved when testing the methods (and reported in the literature) have been linear. It is true that complex nonlinear functions cause difficulties of their own and the characteristics of the solution methods may be disturbed. On the other hand, features concerning nonlinear problems may remain unnoticed with MOLP problems. On the whole, the comparisons available are not of too much help if one is looking for a method for a nonlinear problem, and more contributions in this area are needed. Nevertheless, we review some of the comparisons published.

1.2.2. Noninteractive Tests

An MOLP problem for determining the most economical combination of grape growing and wine production in Hungary is solved by the weighting method, the ε-constraint method, lexicographic ordering and the weighted L_1- and L_2-metrics with normalized objective functions in Szidarovszky and Szenteleki (1987). It is observed that different solutions are obtained with each method. It is also stated that the weighted L_1- and L_2-metrics with normalized objective functions produce the most uniform distribution of objective vectors. Finally, the weighted L_1-metric is seen as the most convenient way for generating Pareto optimal solutions in large-scale MOLP problems.

A linear problem in the mining industry is solved in Peterson (1984) by the weighting method, the ε-constraint method, the method of weighted metrics with and without denominators, and by lexicographic ordering. The solutions obtained from the other methods are utilized in the method of weighted metrics and all the solutions are analysed. The conclusion is that solution methods should be applied so that they complement each other.

1.2.3. Interactive Tests with Human Decision Makers

No interactive methods were included in the comparisons mentioned thus far. The following comparisons involve interactive methods.

It is described in Dyer (1973b) how nine (student) decision makers were presented with an MOLP problem involving choosing an engine for a car. They were first asked to suggest an approach and then compare it with the GDF method and a trial-and-error procedure. In the trial-and-error procedure the decision maker was simply asked to enter an objective vector and the procedure

stated whether it was feasible or not. The decision makers were assumed to explore the feasible objective region until they were unable to find more preferred solutions.

The criteria in the evaluation were the ease of using the method and the confidence in the solution obtained. The results obtained favoured the GDF method. Thus, a conclusion could be drawn that the GDF method can successfully be used by untrained decision makers.

The performance of the GDF method, STEM and the trial-and-error procedure (the same as that used by Dyer) is compared from the point of view of a decision maker in Wallenius (1975). A total of 36 business school students and managers from industry were employed as decision makers. The following aspects of the methods were compared: the decision maker's confidence in the solution obtained, ease of use and understanding of the method, usefulness of the information provided, and rapidity of convergence. The linear management problem to be solved contained three objective functions.

The results are analysed statistically in Wallenius (1975). One interesting conclusion was how well the trial-and-error procedure competed with the more sophisticated methods. Nevertheless, Wallenius points out that its performance might be weakened if the problems were more complex. Difficulties in estimating the marginal rates of substitution weakened the overall performance of the GDF method. Thus, Wallenius suggests that research should be directed to finding ways of better adjusting methods to suit the characteristics of a human decision maker.

The results of Dyer and Wallenius concerning the GDF method differ remarkably. Some trials analysing the reasons are presented in Wallenius (1975).

The capabilities of the ZW, the SWT, the Tchebycheff and the GUESS methods (without the upper and lower bounds) are compared in Buchanan and Daellenbach (1987) from the point of view of the user in solving a linear three-objective optimization problem. The problem concerned the production of the electrical components of lamps. A total of 24 decision makers (students and academic staff) were employed. The criteria in the comparison were partly the same as those used by Wallenius. In addition to confidence in the final solution, ease of use and ease of understanding the logic of the method, CPU and elapsed time were compared. The most important criterion was the relative preference for using each method. The conclusions are that the Tchebycheff method was clearly preferred to the other methods and the ZW method came out the worst in relation to the first four criteria. The SWT method was in the middle. The GUESS method performed surprisingly well. On the basis of this experiment one can say that decision makers seem to prefer solution methods where they can feel that they are in control.

Experimental evaluations of interactive methods with 24 decision makers (students) and two three-objective MOLP problems are reported in Buchanan (1994). The methods involved were the Tchebycheff method, the GUESS method and the simplified interactive multiple objective linear programming

(SIMOLP) method (by Reeves and Franz) based on fitting a hyperplane through a set of extreme point solutions. The criteria of the evaluation were basically the same as in Buchanan and Daellenbach (1987).

A two-stage solution approach was tested where an introductory phase was performed before employing the actual solution methods. In the introductory phase a set of Pareto optimal solutions was generated and the decision makers were asked to provide a rating for each solution. A linear regression model was fitted according to the ratings and the decision maker was asked for feedback as to the correctness of the preference structure formed.

The motivation for using two phases was the following. It has been suggested at times that the opportunity of familiarizing the decision maker with the problem to be solved should improve the actual solution process. However, the experiment did not support this assumption. Either the form of the preference structure to be fitted was incorrect or the ratings were too difficult to provide.

The methods tested had remarkable philosophical differences. SIMOLP was the most structured and GUESS the most unstructured method. In general, the GUESS method was the most favoured of the three. Added to this, the fact that SIMOLP was the least popular method, there is clearly a preference for less structure in the solution method. The SIMOLP method required the decision maker to select the least preferred solution from a given set. This proved out to be difficult. It seemed that it is easier to select the best rather than the worst alternative. (This observation must be related to the nature of the problem to be solved. The opposite may be true in some other cases.)

It is interesting that the Tchebycheff method was rated better than the GUESS method in Buchanan and Daellenbach (1987), whereas GUESS was better than the Tchebycheff method in Buchanan (1994). In the latter test, the GUESS method had been supplemented by allowing the decision maker to specify upper and lower bounds for the objective functions. This may have improved the functionality of the method. Once again, the decision makers liked being in control of the solution process. However, they would have liked to receive pairwise trade-off information to support the process.

The decision makers were also asked to state the solution method they would prefer to use next time. The most important element in the preference proved to be familiarity with the method.

The method of Steuer and the ZW method (in Subsection 5.13.4 of Part II) are compared in Michalowski (1987). Five decision makers from the planning department of a factory were employed. A linear production planning problem with three objective functions was solved and the evaluation criteria were not fixed in advance, although the main interest was in the decision phase. The decision makers had critical comments concerning both the methods, and each of them obtained a different final solution. One can say that the decision processes by the ZW method terminated slightly faster than those by the method of Steuer.

The method of Steuer and STEM are tested in Brockhoff (1985). A total of 147 decision makers were employed to solve six problems involving purchasing cars. The results and progress are analysed according to several criteria, with the method of Steuer emerging with the best outcomes on the average.

An experiment on the differences in the philosophies of methods for continuous compared to discrete problems is presented in Corner and Buchanan (1995, 1997). In Corner and Buchanan (1997), the continuous GUESS method, a modified ZW method and a discrete SMART method (based on constructing a value function) were used to solve a production planning problem with three objective functions by 84 undergraduate students as decision makers. The problem was nonlinear and had continuous variables. The main interest was to determine the ability of the methods to capture the preferences of the decision maker. In other words, how well the methods were able to find desirable solutions and how much the decision makers liked the methods. The time spent on each solution process was also recorded.

One of the conclusions is that the continuous methods were better and faster than the discrete method. The GUESS method was rated easiest to use and to understand. All the methods produced different solutions of which the one generated by the GUESS method was ranked the best. The order of the methods used was found to have no effects on the results. The exception was the case when SMART was used first. Then the solutions obtained with the other methods were statistically the same. In addition, it was observed that a weighted additive value function explaining their ranking behaviour could be found for most decision makers.

Another experimental test involving the ZW method and the GUESS method is reported in Buchanan and Corner (1997). The emphasis was in testing whether any anchoring effect can be explained by the structure of the solution method. A number of 84 students acted as decision makers and solved a nonlinear problem with three objective functions. The conclusion was that an anchoring effect could be seen with the structured ZW method but less so with the free search method GUESS. Thus, it can be deduced that the selection of the starting point is even more crucial with more structured methods than with less structured methods.

Some comparisons of continuous and discrete methods are also presented in Korhonen and Wallenius (1989b). A continuous MOLP problem with five objective functions concerning the allocation of a student's time between study, work and leisure was solved by 65 student decision makers. The five methods compared were all based on the reference direction approach. Only the specification of the reference direction varied. The original way of using aspiration levels was found to be clearly superior to the others.

A more detailed review of the above-described and some other empirical studies involving real decision makers is given in Olson (1992). However, no final conclusions can be drawn from the experiments. The reason is that the test settings and the samples are not similar enough.

1.2.4. Interactive Tests with Value Functions

Next we review some comparisons utilizing value functions to replace decision makers.

An MOLP problem with three objective functions concerning operations planning in the natural gas business is solved by several methods in Mote et al. (1988). A nonlinear value function was employed instead of human decision makers. The problem was solved by the GDF, the SWT and the ZW methods, STEM, goal programming, and the method of Steuer. Only standard LP codes were utilized in the calculations. No single technique was shown to be superior. The methods had differences concerning the burden upon the decision maker and ad hoc and non ad hoc properties.

The method of Steuer and the method of Franz (an interactive adaptation of weighted and lexicographic goal programming) are compared in Gibson et al. (1987) in solving several randomly generated MOLP problems. Different value functions were used to replace the decision maker. The aim of the comparison was to investigate the applicability of the methods to different situations with the help of statistical tests. The number of iterations was also recorded. The conclusion is that, for example, the number of iterations and whether all the objective functions are of relatively equal importance or not are important in the selection of a method. These criteria lead to different recommendations respecting methods.

The Tchebycheff method and the SIMOLP method (by Reeves and Franz) are compared in 15 MOLP test problems with four objective functions using both linear and nonlinear value functions to replace the decision maker in Reeves and Gonzalez (1989). The comparison criteria were the quality of the solution (how far the best solution found was from the best extreme point), user-friendliness, computational requirements, whether nonextreme solutions could be found, number of iterations needed and flexibility. The Tchebycheff method was used in the comparison because of its promising performance in the test reported in Buchanan and Daellenbach (1987). The main difference between the two methods is that the SIMOLP method moves away from the least preferred alternative whereas the Tchebycheff method moves toward the most preferred one. Thus, the SIMOLP method is more flexible and it is easier for the decision maker to change her or his mind. Further, the SIMOLP method needs much less calculation.

The SIMOLP method was able to find slightly better solutions at less computational cost in most problems even with the nonlinear value function. The fact that the SIMOLP method is limited to Pareto optimal extreme points did not seem important in the tests. However, Reeves and Gonzalez (1989) suggest combining the advantages of both methods. Either the decision maker can choose at each iteration which method to utilize for the next iteration, or the flexibility of the SIMOLP method can be used first and the ability of the Tchebycheff method to produce nonextreme solutions can be used in the last iterations.

Another summary of comparisons published is given in Aksoy et al. (1996). The presentation summarizes six comparisons with human decision makers and fourteen studies utilizing value functions. The aspects treated are the comparison criteria, types of decision makers, nature of the test problems to be solved, form of value functions, starting solutions, stopping criteria, and ordering of methods and test problems. The obvious conclusion from this is that there is an urgent need for further comparative evaluations of nonlinear multiobjective optimization methods.

1.2.5. Comparisons Based on Intuition

A characteristic shared by the evaluations to be described next is that they are based on intuition and insight rather than practical experiences and tests. The comparative table in Section 1.1 could equally have been included here.

A collection of the features of five nonlinear interactive methods is presented in Masud and Zheng (1989). The methods are compared with regard to eleven items, for example, the certainty of obtaining a Pareto optimal solution, the optimization technique used, the type of information required from the decision maker, computational complexity compared to the GDF method, and the number of iterations needed with input from the decision maker compared to the GDF method. A similar table comparing the decision maker's burden, ease in actual use, effectiveness and handling of inconsistency is collected in Shin and Ravindran (1991) for ten methods. A further classification and evaluation of methods according to 21 criteria is given in Rietveld (1980).

The number of items a decision maker has to assess simultaneously and per iteration for eight different methods in a medium size linear problem are tabulated in Kok (1985). It is concluded that the method of displaced ideal, the interactive multiple goal programming method and STEM are promising because their presumptions are realistic. In Kok (1986), the learning effects, information load, effort of technical support and group decision capabilities are evaluated for five methods: ZW, interactive multiple goal programming, ISWT, STOM and pairwise-comparisons. No strict preference can be expressed.

A total of 19 interactive methods for MOLP problems are listed according to three characteristics in Larichev et al. (1987). The characteristics are the reliability of the way information is elicited from the decision maker, insignificant sensitivity to random errors on the part of the decision maker and good speed of convergence. The basic principles of the methods are also introduced. In addition, the features of STEM, the GDF and the Tchebycheff methods, the reference point method and the reference direction approach, among others, are tabulated in Vanderpooten and Vincke (1989) and Vincke (1992, p. 105). The criteria are, for instance, assumptions of the existence of a value function, applicability, trial and error support, mathematical convergence, the number of questions posed and the computational burden.

The bi-reference procedure is compared to STEM, the GDF method, the ZW method and the reference direction approach in Michalowski and Szapiro

(1992). The idea is to compare the performance of the bi-reference procedure to the published results of the other methods.

Finally, we mention some other comparative studies. Characteristic values in optimizing the multiobjective layout of a conical shell by the GDF method, STEM and three other methods are reported in Eschenauer et al. (1990b). As far as the relative performance of STOM and the Tchebycheff method is concerned in finding a solution to a linear sausage blending problem in Olson (1993), the main intention is to emphasize the power of the weighted Tchebycheff metric in multiobjective optimization.

1.3. Selecting a Method

Choosing an appropriate solution method for a certain multiobjective optimization problem is not easy, as has been made abundantly clear. None of the existing methods can be labelled as the best for every situation, since there is a multiplicity of aspects to consider and many of the comparison criteria are of a somewhat fuzzy character. The features of the problem to be solved and the capabilities and the type of the decision maker have to be charted before a solution method can be chosen. Some methods may suit some problems and some decision makers better than others. Let us sum up by offering some general guidelines and a decision tree.

1.3.1. General Guidelines

Several different comparison criteria were already mentioned in Section 1.2 in connection with the tests reported. Some of the criteria to consider when evaluating methods are also collected in Hobbs (1986). These selection criteria are appropriateness, ease of use, validity and the sensitivity of the results to the choice of method. Appropriateness means that the method is appropriate to the problem to be solved, to the people who are to use it and to the institutional setting where it is to be implemented. Ease of use refers to the effort and the knowledge required from the analyst and the decision maker. Validity means that the method measures what it is supposed to and the assumptions set are consistent with reality. The sensitivity of the results to the choice of method expresses the desire that solutions obtained by the method do not significantly differ from those of other methods. If the method chosen has a significant effect on decisions, then the relative validity of different methods should be considered. If the form of the method does not matter, then the most important criteria are ease of use and appropriateness.

The number of crucial criteria in selecting a solution method is reduced to three in Stewart (1992). The input required from the decision maker must be meaningful and unequivocal, the method must be as transparent as possible and it must be simple and efficient.

The role of the decision maker is important and should be taken seriously. Many experiments have shown that decision makers prefer simpler methods because they can more easily understand such methods and they feel more in control. The valuation placed on some methods may increase if the decision makers can practice using them or obtain advice. An important fact to keep in mind is that theoretically irrelevant aspects, such as question phrasing, may affect the confidence that the decision maker feels in the method. The concept of the decision maker's confidence is analysed further in Bischoff (1986).

Other important criteria for the decision maker in selecting the solution method are, for example, the simplicity of the concepts involved, possibilities of interaction, the ease with which the results can be interpreted and the chances of choosing the most preferred solution from a wide enough set of alternatives. The method must also fit the decision maker's way of thinking. The language of communication between the decision maker and the method (solution system) must be understandable to the decision maker. (S)he wants also to see that the information (s)he provides has a (desirable) effect on the solutions obtained.

One more element, not mentioned thus far, in the selection of a method is how well the decision maker knows the problem to be solved. If (s)he does not know its limitations, possibilities and potentialities well, (s)he needs a method that can provide support in getting acquainted with the problem. In the opposite case, a method that makes it possible to focus directly on some interesting sector is advisable. Ways of identifying appropriate methods for different types of decision makers are needed.

1.3.2. Method Selection Tools

Few universally applicable guidelines have been given for the method selection problem in the literature. Let us mention some of them including even approaches for discrete problems.

An attempt to assist in the selection of a solution method is presented in Gershon and Duckstein (1983). The selection problem is modelled as a multiobjective optimization problem. A set of 28 criteria for the selection are suggested and they are divided into four groups. Only the criteria in the last group have to be considered every time the selection algorithm is applied. The criteria take into account the characteristics of the problem, the decision maker and the methods. Many types of problems are taken into consideration in the criteria (e.g., discrete and continuous variables). The model contains 13 solution methods from which to select. The set of methods can naturally be modified. The number of selection criteria can also be varied to include only those relevant to the problem to be solved. Finally, after the methods have been evaluated according to the selection criteria, the resulting multiobjective optimization problem is solved by the method of the global criterion (e.g., L_1-metric).

A related procedure is suggested in Tecle and Duckstein (1992). There, a set of 15 methods is evaluated with respect to 24 criteria in four classes. The weighted L_p-metric is used in each class and another weighted L_p-metric is used

to combine the classes and obtain the best method. For the example problem provided, the weighted L_p-metric turns out to be the best method. One may wonder whether the weighted L_p-metric favours itself or whether this is a mere coincidence. Some critique of the approach is also expressed in Romero (1997).

Different decision trees and rules for providing assistance in selecting a method for multiattribute decision analysis problems are described in Hwang and Yoon (1981) and Teghem et al. (1989). However, as criticized in Ozernoy (1992a), to design a comprehensive and versatile decision tree usually results in an explosion in the number of nodes. Another problem with decision tree diagrams is what to do when the user answers 'I do not know.'

An expert system for advising in the selection of solution methods for problems with discrete alternatives in proposed in Jelassi and Ozernoy (1989). Steps in the development of another expert system for selecting the most appropriate method for discrete problems are described in Ozernoy (1992a, b). The questions posed by the system are based on if/then rules. They lead to recommending a method or stating that no method can be recommended. The user of the system can also always ask why a particular question is posed.

1.3.3. Decision Tree

Little advice exists for selecting a method for nonlinear and continuous problems. Therefore, despite the above-mentioned pitfalls and faults in decision-tree diagrams, we nevertheless present one in Figure 1.3.1. The tree has primarily been created on the basis of plain theoretical facts concerning the assumptions imposed by the methods on the problem to be solved and secondarily according to the preferences of the decision maker. Because of space limitations it has not been possible to include all the properties.

The decision tree includes the twelve interactive methods described in Part II. Only those methods are included that have been presented in more detail or whose main features have been introduced. Remember that in practice, the functioning of a method may not always require that all its technical assumptions are satisfied (as stated, for example, in Zionts (1997a, b)). Or it may even be impossible to verify all the assumptions. If some of the assumptions are not valid, some of the results may be incorrect, but this does not necessarily mean that the method will not work in some contexts. The results may still be adequate for practical purposes. This must be kept in mind when studying the decision tree. Nonetheless, the assumptions provide some guidelines to follow.

The starting node is situated on the left. The tree diagram has been created in such a way that only the answers 'yes' or 'no' are possible. Whenever the immediate answer is 'I do not know,' the answer 'no' can be given. In order to avoid confusing the picture any further, the words yes and no have been replaced by arrows of different types. Continuous lines represent positive answers and broken lines stand for negative ones. In addition, 'no' arrows always leave a node to the right of the 'yes' arrows.

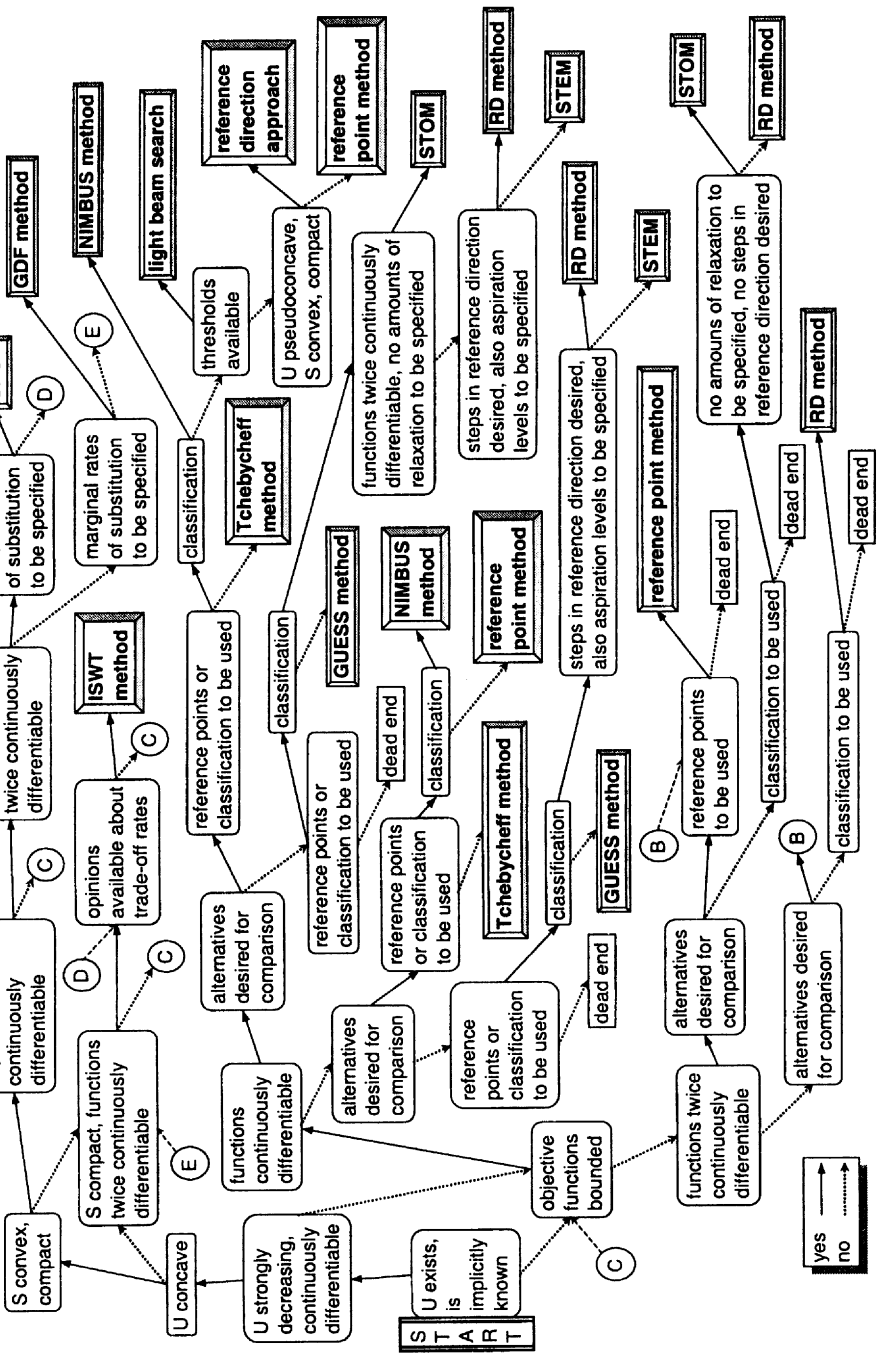

Figure 1.3.1. Tree diagram.

The nodes containing only capital letters are used in two different cases. The first is to avoid repetition. In the second, no method can be found along the path followed. In that case, one can try another path. The aim has been to allow as many previous answers as possible to be exploited. Thus, some dead ends may be avoided.

The same method may be reached by following different paths. In this case varying questions may be needed in order to separate the methods. That is why either general or more detailed questions leading to the same method are used.

As already repeated several times, selecting the solution method is a difficult and important task. After describing each method (see Part II), we have tried to indicate what it has in its favour and what its drawbacks are. These matters, of course, are always more or less subjective.

2. SOFTWARE

The development of computers and the improvement in the speed, storage capacities and flexibility of computing facilities have made it possible to produce more sophisticated and demanding software for solving multiobjective optimization problems. Efficient computers enable, for example, the implementation of interactive algorithms, since they can produce sufficiently fast responses for the decision maker without the user getting frustrated waiting.

Nevertheless, taking into account the multiplicity of methods developed for solving nonlinear multiobjective optimization problems, the number of widely tested and user-friendly computer programs that are generally available is small. At least they are difficult to find. Most implementations are done for academic testing purposes and their existence is not advertised. In other words, there is a real need for functional and reliable software for solving nonlinear multiobjective optimization problems.

2.1. Introduction

Most of the software packages developed for multiobjective optimization problems can be termed multiobjective decision support systems, and they form one class of decision support systems. *Decision support systems* (DSSs) can be defined as interactive computer-based systems designed for helping and assisting in the decision-making process. Their main objects are to help decision makers in solving problems more efficiently and making better decisions.

The main components of a decision support system are a model, an optimizer (solver) and an interface between the model, the optimizer and the user. By an interface we mean the input language and style, exchange of information and presentation of the results. It should be remembered that the human-computer interface must be designed with at least as much care and effort as the other components of the system.

Another, similar characterization of a decision support system is that it consists of a model, by which is meant a mathematical algorithm, data management and interface. It is essential that the model reflects the preference structure of the user. That is why it is important that users take an active role in developing decision support systems.

K. Miettinen, *Nonlinear Multiobjective Optimization*

The role and the requirements of the model, the optimizer and the interface in the multiobjective optimization environment are outlined, for example, in Jelassi et al. (1985). It is useful to have capabilities of self-learning and model updating in a decision support system. The interface is an important factor in relation to the user-friendliness of the system.

One can state that developing software for multiobjective optimization problems is once again a multiobjective optimization problem in itself, and proper planning is essential. Several (conflicting) objectives to be taken into consideration in multiobjective software design and realization are collected in Olkucu (1989). Among them are a short development time, long product life, easy and cheap maintenance, reliable implementation of the algorithm, an efficient user interface and a large number of potential users. Of no minor importance in this regard are the selection of the realization environment (including the operating system) and development tools.

Features to be taken into consideration when designing decision support systems are also handled in Lewandowski (1986). Different definitions of user-friendliness and rules for dialogue design are given.

It should be pointed out that while a great deal of effort has gone into developing the methodological and computational aspects of the systems, the interface between the system and its user is often of poor quality. This is a serious weakness, since no matter how brilliant the methodology and its implementation are, it will be discarded if the interface does not suit the user. In any case, the algorithms must be implemented in such a manner that computer-technical requirements do not overshadow the real problem and non-skilled persons can also use the programs. One way to try to improve the situation is to provide different interface possibilities for the same system for computer specialists, trained users and average users.

An effort of measuring the effectiveness of decision support systems is described in Sainfort et al. (1990). Even though it is widely assumed that decision support systems really do help in decision making and problem solving, research results in this important area are few. Added to this is the fact that there is currently no general theory about problem solving because of its complexity. Group decision support systems are mostly handled in Sainfort et al. (1990), but the conclusions favouring decision support systems are general. It is demonstrated that decision support systems increase the understanding of the problem, reduce frustration in the problem solving and contribute to progress in the solution process.

To put if briefly, a decision support system should be easy to use, it should capture the thinking procedure of the decision maker, it should support different decision styles and it should help the decision maker to structure different situations. Other desirable characteristics of decision support systems are listed in Weistroffer and Narula (1997).

The user of the decision support system may need guidance and training to be able fully to make the most of it. Such a step may increase usability and render the system more user-friendly than before.

An interesting question is raised in Verkama and Heiskanen (1996) about how research concerning both methodology and decision support system software should be reported in the literature. Verkama and Heiskanen suggest that numerical examples should accompany both the algorithms proposed and the software described. This would enable any interested reader to use the example to understand details not presented in the paper.

2.2. Review

Existing software packages up to the year 1980 are listed in Hwang et al. (1980). These programs were mainly developed for linear and goal programming problems. They are rather primitive when compared with modern computer facilities. A somewhat more up-to-date list of decision support systems developed to aid in multiobjective optimization and multiattribute decision analysis problems up to the year 1988 is collected in Eom (1989). However, the presentation is only cursory. In addition, a classification of the system applications is provided. Some software implementations are also mentioned in Weistroffer and Narula (1991), whereas the overview in Buede (1996) handles software for discrete problems only.

The latest state of decision support systems for multiple criteria decision-making problems up to the year 1997 is presented in Weistroffer and Narula (1997). Systems for both continuous and discrete problems are listed with information about where they can be obtained. Unfortunately, from among the seventeen continuous products mentioned only six are applicable to general nonlinear multiobjective optimization problems. Of these GRS generates and illustrates the Pareto optimal set implementing the generalized reachable sets method (see Section 3.6 in Part II). The actual solvers are CAMOS, DIDAS, LBS, MONP-16 and NIMBUS, where LPS is an implementation of the light beam search (see Section 5.9 of Part II). The others have already been introduced in Part II except for CAMOS (see the end of this section). DIDAS, the implementation of the reference point method, was described in Subsection 5.6.4, MONP-16 implementing STOM was mentioned in Subsection 5.8.4 and implementations of NIMBUS were handled in Subsection 5.12.8.

The current state of the software development can be inquired from A. Lotov, Russia (GRS), A. Osyczka, Poland/Japan (CAMOS), J. Granat, Poland (DIDAS), A. Jaszkiewicz, Poland (LBS), L. Kirilov, Bulgaria (MONP-16) and K. Miettinen, Finland (NIMBUS). The WWW-NIMBUS system is available at http://nimbus.math.jyu.fi/.

Software for discrete problems is more easily available than software for continuous problems. For example, several discrete systems have been developed

into commercial products. Yet, as emphasized in Buede (1996), even those software developers concentrate too closely on features of analysis at the expense of user-friendliness, as mentioned earlier.

Among software products for solving MOLP problems is VIG by P. Korhonen, Finland (see Subsection 5.10.3 in Part II). Let us also mention a package of subroutines, called ADBASE, by R.E. Steuer, USA (see Steuer (1986, pp. 254–267)). ADBASE contains, for example, tools for generating Pareto optimal extreme points. These are examples of the generally available products for linear problems.

The situation is worse for continuous nonlinear problems. Most of the software implementing the extensive amount of existing multiobjective optimization methods is neither commonly available nor widely known.

One explanation sometimes mentioned is the lack of a free and reliable nonlinear solver that could be integrated and distributed with the software. Most software products have been implemented for academic testing purposes and have not been updated along with the development of computer facilities. Consequently, their existence is not advertised. Simply designing and realizing a functional user interface is demanding. One must assume that the need has not been large enough to motivate the work, or the need has not been realized because good solution tools have not been available.

However, some implementations were mentioned in connection with the method descriptions in Part II. No detailed information was given, since the implementations are under continuous development and the details may be out-of-date at any moment.

Software comparisons reported in the literature mainly concern programs for multiattribute decision analysis. We simply mention that seven microcomputer implementations are presented and compared in Colson and de Bruyn (1987). Five of them are intended for multiattribute decision analysis. An implementation of STEM is also reported. In addition, the main features and requirements of eight microcomputer software packages are introduced in Lotfi and Teich (1989). One of them is VIG and the other seven are for discrete alternatives.

There exist several software packages for general single objective optimization problems that also contain some possibilities for noninteractive multiobjective optimization. Let us briefly indicate some of them. The implementation of the MPB method (see Section 2.2 in Part II), called MPBNGC by M. Mäkelä, Finland, is among them.

CAMOS (Computer Aided Multicriterion Optimization System) has been developed to treat especially nonlinear computer aided optimal design problems (see Osyczka (1989b, 1992) and for an earlier version Osyczka (1984)). CAMOS produces Pareto optimal solutions with different generating methods.

The methods for identifying (weakly) Pareto optimal solutions are the weighting method with or without normalizing the objective function, the ε-constraint method, the method of the global criterion and the method of

weighted Tchebycheff metric. Problem (2.1.4) of Part II is also used. For more details, see, for example, Osyczka (1984, 1992). Different underlying single objective optimization algorithms may be used.

The functioning of CAMOS is illustrated by two practical problems in Osyczka (1992, pp. 93–125). They are the optimal design of multiple clutch brakes and the optimal counterweight balancing of robot arms.

NOA, a collection of subroutines for minimizing nondifferentiable functions subject to linear and nonlinear (nondifferentiable) constraints, is described in Kiwiel and Stachurski (1989). NOA is applicable to multiobjective optimization problems since the single objective function to be minimized is assumed to be a maximum of several functions. Thus, for example, some achievement functions can be optimized.

Let us finally mention the optimization toolbox of the MATLAB system including the weighting method, the ε-constraint method and a modification of goal programming. Naturally, other multiobjective optimization algorithms may be coded within the MATLAB environment, taking advantage of the powerful single objective solvers and graphics available.

3. GRAPHICAL ILLUSTRATION

3.1. Introduction

Graphical illustration plays an essential role when designing modern software user interfaces. Graphics may be used to describe the problem, to assist the decision maker in specifying values for problem parameters or to illustrate the contents and the meaning of questions posed by the algorithms. In such realizations, the upper limit lies in one's imagination.

In spite of the more general possibilities, we restrict our treatment in this chapter. By graphical illustration we here mean the ways of presenting several alternative objective vectors to the decision maker. To be convinced of the need for such illustration one has only to examine the interactive methods described in Part II. Good graphical illustration helps the decision maker to gain a better insight into the problem and the different alternatives generated.

As computers have developed, more attention has been paid towards the role and the possibilities of computer graphics in building human-computer interfaces. Nevertheless, utilizing graphical illustration does not mean that the limits on human information processing capacity are transcended. This means that there is no sense in trying to offer too many objective vectors for evaluation, no matter how clear the illustrations are.

Several psychological tests are summarized in Miller (1956) to prove that the span of absolute judgment and the span of immediate memory in human beings is rather limited. We cannot receive, process or remember large amounts of information. The magical number seven plus or minus two appears in several tests and in several ways. However, no number can be regarded as an absolute limit. Everything depends on the circumstances. Still, the findings of Miller are to be kept in mind when deciding the number of alternatives to be presented to the decision maker or the number of objective functions to be treated (if these can be affected). Miller's findings must also be remembered when expecting exact information from the decision maker. Let us mention that, for example, seven ways of decreasing the number of alternatives are presented in Graves et al. (1992).

Naturally, many different ways for illustrating objective vectors can be thought of. However, elegant graphics must not be an end in itself. The graphics must be easy to comprehend by the decision maker. On the one hand, not

K. Miettinen, *Nonlinear Multiobjective Optimization*

too much information should be allowed to be lost and, on the other hand, no extra unintentional information should be included in the presentation.

3.2. Illustrating the Pareto Optimal Set

In the case of two objective functions, graphical illustration of the objective space is effective. The feasible objective region and, especially, its Pareto optimal subspace can be sketched on a plane. If this is not possible, the available objective vectors can be plotted in the objective space. As far as three objective functions are concerned, the Pareto optimal set can be expressed by three projections on a plane, as suggested, for example, in Meisel (1973). However, the interpretation of such information is far more difficult for the decision maker.

Another way of illustrating the Pareto optimal set of three objective functions is to draw a two-dimensional plot with fixed values assigned to the third objective function. There is a resemblance here with topographic maps. Such an approach is handled in Bushenkov et al. (1995) and Lotov et al. (1997), where so-called decision maps are used. Several level sets of the third objective function are drawn in the picture of the Pareto optimal hull of the first two objective functions. These sets are called efficiency frontiers. If there are more than three objective functions, several different pictures can be drawn each having fixed values for the other objective functions. For example, in the case of five objective functions a matrix of decision maps may be displayed. There, the fourth objective function has the same fixed value in every picture in each row and the fifth objective function has the same fixed value in every picture in each column. In addition, scroll-bars and animations can be used. According to its developers, this approach works for up till seven objective functions.

3.3. Illustrating a Set of Alternatives

Below, we present some ways of illustrating a set of alternative objective vectors graphically. Some of the ways are clarified by applying them to an example of three alternative objective vectors of a problem with three objective functions.

3.3.1. Value Path

A widely used way of representing sets of objective vectors is to use *value paths*, as suggested, for example, in Geoffrion et al. (1972) and Schilling et al. (1983). This means that horizontal lines of different colours or of different line styles represent the values of the objective functions at different alternatives. In other words, one line displays one alternative. This is depicted in Figure 3.3.1. The bars in the figure show the ranges of the objective functions

in the Pareto optimal set. If the ranges are known, they give additional information about the possibilities and limitations of the objective functions. Note that each objective function can have a scale of its own. Examples are suggested in Törn (1983) of how to display the scales of the objective functions.

Value paths are a recommendable method of illustration because they are easy to interpret. For example, it is easy to distinguish non-Pareto optimal alternatives if they are included. Further, even a large number of objective functions or alternatives causes no problems. Value paths are used, for example, in WWW-NIMBUS (see Subsection 5.12.8 of Part II) and the visual interactive sensitivity analysis system VISA, see Belton and Vickers (1990).

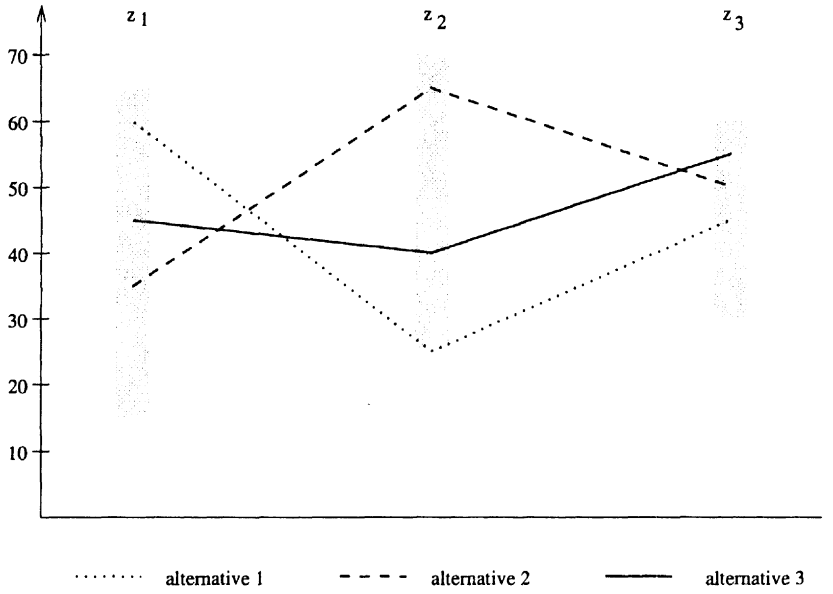

Figure 3.3.1. Value paths.

In value path illustrations the roles of the lines and the bars can also be interchanged. Then bars denote alternatives and lines represent objective functions. In this case, possible different scales of the objective functions have to be interpreted differently (see, e.g., Hwang and Masud (1979, p. 109)). This reversal of roles has been utilized, for instance, in the first implementations of the reference direction approach (described in Section 5.10 of Part II), and its counterpart for discrete problems, called VIMDA, see Korhonen (1986, 1991a). The idea in VIMDA is that when the user horizontally moves the cursor to a bar representing an alternative, the corresponding numerical objective values are presented.

3.3.2. Bar Chart

Value paths are an effective means of presenting information without overloading the decision maker. Another general mode of illustration is to use *bar charts*. This means that a group of bars represents the alternative values of a single objective function, as in Figure 3.3.2. The bars of the same colour indicate one alternative. Separate ranges for objective functions are possible as well. Parallel ideas have been realized, for example, in DIDAS and WWW-NIMBUS, treated in Subsections 5.6.4 and 5.12.8 of Part II, respectively.

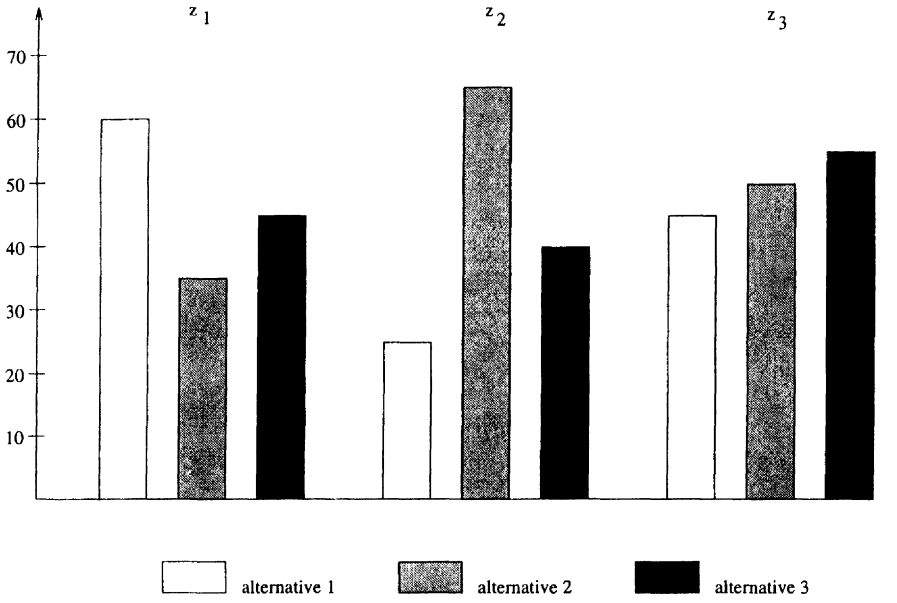

Figure 3.3.2. Bar chart.

Naturally the roles of the alternatives and the objective functions can be interchanged in bar charts as well as in value paths. This, of course, means that the order of the bars is altered. This is possible, for example, in WWW-NIMBUS.

An alternative to using separate ranges for the objective functions is to provide bar charts and value paths using both absolute and relative scales. This is advisable in particular if the ranges of the objective functions vary widely. This option is also available in WWW-NIMBUS.

3.3.3. Star Coordinate System

It is suggested in Mañas (1982) that objective vectors can be represented in a *star coordinate system*. For example, an alternative of three objective functions is represented as an irregular triangle. This requires the ideal objective vector and the (possibly approximated) nadir objective vector to be known. An example is given in Figure 3.3.3. Each circle represents one alternative objective vector. The ideal objective value is at the centre and the component of the nadir objective vector is at the circumference. Each ray represents one objective function. The area of the star depicts each alternative. See details in Mañas (1982).

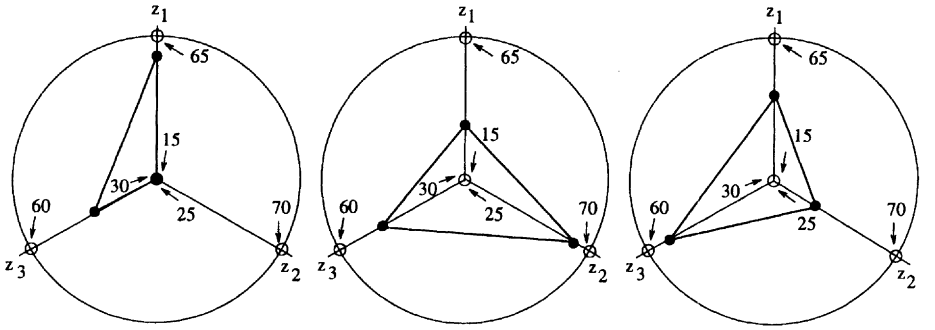

Figure 3.3.3. Star coordinate system.

One can say that in the star coordinate system an alternative is better the smaller the area of the star. If the order of the objective functions is altered, the shape of the star changes. This can be considered a weakness of the system, as stated in Tan and Fraser (1998). Tan and Fraser also suggest a modified star graph to include the weight information of the decision maker (if available) in the same display with the objective values.

3.3.4. Spider-Web Chart

Ideas similar to the star coordinate system are exploited in Kasanen et al. (1991). An example is presented in Figure 3.3.4. This form of illustration can be called a *spider-web* because of its shape. Sometimes it is also called a radar chart. Each apex represents one objective function. The outer triangle shows the (possibly approximated) nadir objective vector, the inner triangle (the darkest one) stands for the ideal objective vector and the middle triangle (the grey one) presents one alternative objective vector. Thus, only the middle triangle is different in the different alternatives. These ideas are further developed in Kasanen et al. (1991).

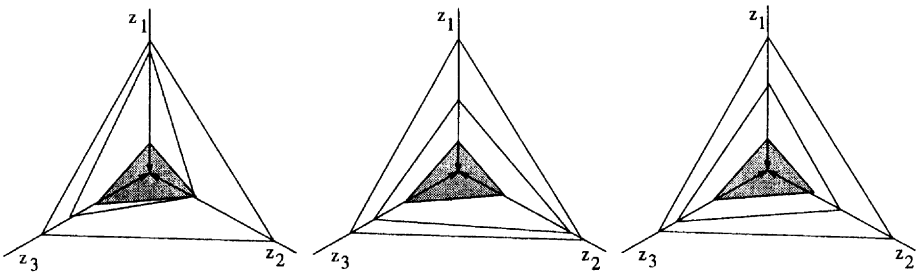

Figure 3.3.4. Spider-web chart.

3.3.5. Petal Diagram

Somewhat parallel ideas to the two previous representations are utilized in Angehrn (1990a, b) when illustrating discrete alternatives in a program called Triple C (Circular Criteria Comparison). A circle is divided into k (the number of objective functions) equal sectors. The size (radius) of each slice indicates the magnitude of the objective value. Here we have one circle for each alternative objective vector. The same idea is suggested in Tan and Fraser (1998) and it is called a *petal diagram*. Each segment of the diagram, that is, each objective function can be associated with a different colour, as in Figure 3.3.5. Notice that the order of the objective functions has no effect on the shape of the diagram. The relations of the different segments are clearly shown. A way of connecting weighting information in the petal figures is suggested in Tan and Fraser (1998). In this case the segments are not of equal size but reflect the weighting coefficients.

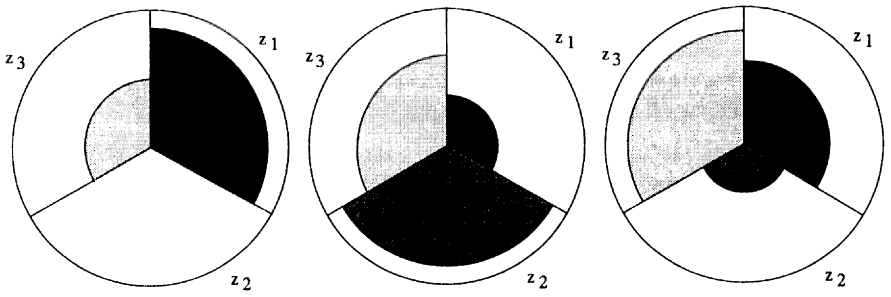

Figure 3.3.5. Petal diagram.

It is mainly a matter of taste in the star coordinate system, the spider-web chart and the petal diagram how the ideal objective vector is situated. One may think that when minimizing the objective functions it is logical to have the ideal area as small as possible. However, the roles can be interchanged so that the ideal objective value is located on the circumference and the nadir

objective value at the centre. In this case the larger the area the better. If this is the setting, the ideal objective values can be replaced with, for example, average objective values. This means that the figures can extend beyond the circumference, stressing values better than the average.

3.3.6. Scatterplot Matrix

The *scatterplot matrix* described in Cleveland (1994) can be adapted for visualizing different alternatives. The scatterplot matrix consists of panels each representing one objective function pair. The dimension of the square matrix is the number of objective functions. Different alternatives can be denoted by different symbols or colours. As can be seen in Figure 3.3.6, each pair is graphed twice with the scales interchanged. This means that either the lower or the upper triangle could be dropper without losing any information. However, displaying the whole matrix makes it easier to compare the objective function values. One can measure the performance of one objective function against the other objectives by having a look at one column or one row. Each objective function can naturally have a range of its own in the panels, as in Figure 3.3.6.

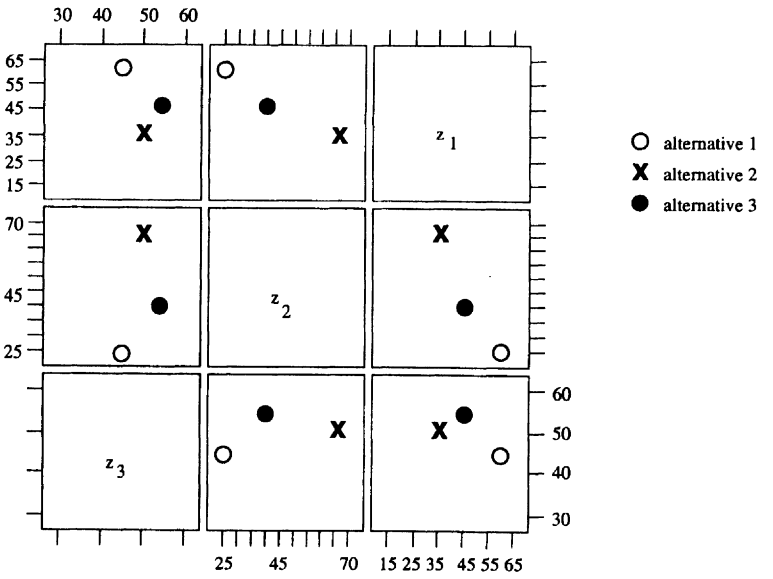

Figure 3.3.6. Scatterplot matrix.

3.3.7. Other Illustrative Means

A graphical display system called GRADS is introduced in Klimberg (1992). GRADS is dynamic and can be applied to problems with about five to twelve objective functions. The decision maker is first asked to indicate two objective functions whose values in the different alternatives are drawn as points in a plane. In this space, the adjacent alternatives are connected with lines. In other words, we have one value path. The decision maker obtains information about the other objective values for one alternative at a time by indicating that point with a mouse. Then, the other objective values are depicted as lines originating from the point considered. The lengths of the lines are proportional to the objective values. The end points of the lines are connected, thus forming triangles of a different colour. The percentage achievements of the alternative in question are also displayed. They are calculated as the difference between the nadir objective value and the current objective value divided by the range. The decision maker can change the alternative considered and the two objectives whose value paths form the base of the display.

Different ideas of graphical illustration are also handled in Korhonen (1991b). One of the ideas is Chernoff's faces, originally developed to illustrate numerical information. The idea is to represent the values of up to 18 objective functions as the characteristics of a face. In other words, the values of each objective function are parametrized to represent some feature of an icon. The icon used must be such that the user can see the icon becoming 'better' as the value of the objective function improves. This is why concepts like symmetry and harmony are important. An icon that people have been used to seeing in a harmonious and symmetrical form is a house. Thus, Korhonen suggests so-called *harmonious houses* to be used as icons. Objective functions are associated with the corner points of the house, the door, the windows or the roof. The aim is that when the values of the objective functions are close to the ideal objective vector, the house is quite harmonious and symmetrical. This type of illustration has especially been intended for pairwise comparison.

Literature describing the graphic presentation of data is summarized in Lewandowski and Granat (1991). It can be concluded that the research done does not provide clear answers regarding what types of data presentation to favour in the decision-making context. Lewandowski and Granat suggest a technique called BIPLOT for the graphical presentation of matrices of rank 2. The set of Pareto optimal objective vectors forms a matrix. This matrix is factorized into a product of two matrices. The vectors in the two matrices are of order two and can be plotted on a plane giving a representation of the original objective vectors. Dynamic BIPLOT in aspiration level-based decision support systems is also described. Another question is how much experience one must have to be able to interpret representations like these.

Several tools for use in creating illustrations are also summarized in Klimberg (1992). One of them is to transform objective vectors into two-dimensional

curves with the aid of Fourier series. In this way all the vectors can be plotted on the same coordinate system for comparison.

Other proposals for the graphical illustration of alternatives are given, for example, in Vetschera (1992). They are based on indifference regions and linear underlying value functions.

Let us finally mention a projection idea called GAIA (Geometrical Analysis for Interactive Aid). It is a part of the discrete multiattribute decision analysis method PROMETHEE and it is described, for example, in Brans and Mareschal (1990) and Mareschal and Brans (1988). The objective functions are first modified to include some preference information of the decision maker and then normalized. These objective functions have some benefits when compared to the original ones. Namely, they are in the same scales, big differences in the objective values are emphasized and small differences are lessened.

Principal component analysis is used in order to find a plane (two dimensions) in which the new objective functions can be projected. The idea is to lose as little information and variation as possible. In other words, the two largest principal components are selected to form the projection plane. The weakness here is that if the objective functions have nonlinear relations, principal component analysis cannot find it.

If selecting the plane is managed well enough, the relations between the new objective functions and the alternative solutions can be seen in their projections. Objective functions are depicted as vectors and alternatives as points on the plane. For example, if two objective functions are highly conflicting, their vectors go in opposite directions, whereas independent objective functions are orthogonal and similar objective functions are oriented approximately in the same direction. From the location of the alternatives one can see how well they perform with respect to each objective function, that is, how near or far they are from each other.

It seems that this GAIA plan ideology is a rather clear method of illustration. However, it has two main limitation. Firstly, the plane contains only a part of the information available. Secondly, the conflict characteristics of the objective functions are not absolute but depend on the alternatives considered.

3.3.8. General Remarks

The problem of how we can determine a priori whether the graphical formats used will aid rather than hinder decision making is examined in Jarvenpaa (1989) by comparative studies. The conclusion is that knowledge concerning the relationship between the presentation format and the decision strategy can facilitate the selection of the presentation format. Special attention is given to the benefits of bar charts and grouped bar charts.

Similar matters are handled in connection with visual interactive simulation in Bell and O'Keefe (1995). For example, it is concluded that the use of visual displays generates solutions that are demonstrably better than those that make

limited use of such displays. This means that different levels of usage of specific displays have an impact on the quality of the solutions generated. In the experiments, bar charts were the most favoured visual displays.

Several existing studies on the applicability of graphs versus tables are analysed in Vessey (1991). According to the theory developed, it is concluded that tables perform better in information acquisition tasks in both time and accuracy of performance. Thus tables are in order when specific data values must be extracted, since they represent discrete data values. If information must be viewed at a glance, evaluated or relationships in the data are of interest, graphs are recommendable. Thus, graphs and tables emphasize different characteristics of the same data.

Using colours in illustrations has advantages and disadvantages. Above all, the colours must be easy to discriminate. Another important issue is that some colours may have specific connotations to the user. Such colours should be avoided as far as possible.

An experimental evaluation of graphical and colour-enhanced information presentation is given in Benbasat and Dexter (1985). Colours improve the readability and understandability of both symbolic and graphical displays. Colours make it easier for the decision maker to associate visually information belonging to the same context or unit since such data are coded in the same colour. Encouraging results with multi-colour reports are mentioned by Benbasat and Dexter (1985), who also stress that tabular representation is the best when a simple retrieval of data is important and a graphical representation is the best when relationships among the data have to be shown. Graphs are visually appealing but sometimes tables are easier to read since they provide exact values.

A recommended way of presenting information to the decision maker is to offer the same data in different forms. In this way, the decision maker can choose the most illustrative and informative representations. The illustrations may also supplement each other. The decision maker can change her or his attention from one figure to another and possibly skip undesirable alternatives before making the final selection. A simple tabular format may be one of the figures. Corresponding ideas are suggested, for instance, in Silverman et al. (1985) and Steuer (1986, pp. 520–522).

An interesting alternative to graphics and numerical values is suggested in Matos and Borges (1997). The idea is to illustrate alternatives in natural language phrases. In this approach, fuzzy membership functions are formed for every objective function defining fuzzy bounds, for example, 'very little', 'little', 'medium', 'very' and 'most' values. An example of alternatives in a washing machine selection problem could be 'most cheap, medium power saver and little fast.' The decision maker is asked for some descriptive information as the basis of the membership functions before the solution process. This is a promising way of illustrating data and it deserves further development.

Finally, one must concede that where a great number of alternatives coexists, the decision maker may get confused no matter how the alternatives are illustrated. In this case, statistical tools, for example, principal component analysis, may be useful.

4. FUTURE DIRECTIONS

In this chapter, we outline some challenging topics for the future development of multiobjective optimization, mainly from a mathematical point of view. In addition, we give examples of promising ideas for research where the first steps have been taken but further work is needed. All the issues mentioned and many others merit further research and examination.

Multiobjective optimization is important, and improved solution methods can bring about change in many areas and aspects of life. Even though multiobjective optimization methods have been applied to solving a variety of problems in many areas of life, such as design problems in engineering, production problems in economics, and environmental control problems in ecology, there continue to exist many new problem types which could benefit highly from multiobjective optimization. Particularly challenging in this respect are real-life problems. There is clearly a need for more contributions reporting on practical applications (making good use of more developed methods).

One interesting type of problems is so-called multidisciplinary re-engineering. It means that old engineering problems, for example, in optimal design, whose solutions have been revised one feature at a time over the course of years, are solved again from the very beginning, taking various aspirations and aspects into consideration at the same time. Obviously this requires tools of multiobjective optimization.

An important challenge for the developers of interactive methods and approaches is how to approach the decision maker. For example, a real experiment with problem-related decision makers in Hobbs et al. (1992) shows that the decision makers were sceptical of the value of multiobjective optimization methods and they in some cases preferred unaided decision making. This means that the methods should not only be user-friendly but also of real help to decision makers. Combining knowledge from the behavioural sciences with method development could usefully serve in this direction.

The methodology of multiobjective optimization must be improved. This means, for example, creating computationally efficient ways of generating trade-off information for more general problem types under less restricting assumptions than those employed thus far. Another aspect is the structure of the methods. On the one hand, providing the decision maker with the opportunity for free search is important. On the other hand, guidance and support must

K. Miettinen, *Nonlinear Multiobjective Optimization*

be available, if desired. This necessitates developing mechanisms for dealing with inconsistencies. In addition, ways of identifying appropriate methods for different problems and different types of decision makers are certainly needed.

It is important not only to develop general methods but also to create algorithms specifically tuned to certain problem types and areas of application. An example is the monograph by Janssen (1992), where methods and decision support tools for environmental management are dealt with.

An alternative to creating new methods is to use different methods in different phases of the solution process. In this way, the positive features of various methods can be exploited to their best advantage in appropriate phases of the solution process. In addition, it may be possible to overcome some of the weaknesses of the existing methods.

An example of the combination of several methods is a meta algorithm endeavouring at consolidating different methods of multiobjective optimization. This is proposed in Steuer and Whisman (1986). The idea is that the same meta program can be transformed into different methods by varying its control parameters. The GDF, the Tchebycheff and the reference point methods with the reference direction approach, STEM, the ε-constraint method and two interactive versions of the weighting method are available. This idea is further developed in Steuer and Gardiner (1990). An important fact to consider, when switching from one method to another in the middle of the solution process, that is, how to maintain the convergence properties, needs further investigation. In Gardiner and Steuer (1994a, b), the meta algorithm is extended into a unified algorithm containing thirteen different interactive methods. A vital element of the algorithm is a matrix describing what kinds of switches are allowed between the methods.

Similar ideas of combining several methods are proposed in Clímaco and Antunes (1991). The system (only for MOLP problems) contains, for example, the ZW method, STEM and VIG. Only problems with three objective functions can be handled. The system has also been implemented. A further developed implementation of the above-mentioned ideas is described in Antunes et al. (1992a) and Clímaco and Antunes (1994). The method base package has been named TOMMIX. Suitable means to support the decision maker in deciding when and how to change from one method to another have still to be explicated. TOMMIX is further extended into SOMMIX for more than three (linear) objective functions in Clímaco et al. (1997).

Another approach to be elaborated is combining methods for continuous and discrete problems. It may, for example, be that a set of solutions is generated for the continuous problem and then ranked by means of discrete methods. Examples of this are presented in Bard (1986), Kok and Lootsma (1985) and Slowinski (1991). One example of these methods, the light beam search, was described in Section 5.9 of Part II.

One can also combine methods of global optimization with multiobjective optimization methods. In this way, one can aim at being able to handle globally

Pareto optimal solutions, instead of locally Pareto optimal ones, in nonconvex problems as well. Ideas of global multiobjective optimization on the basis of clustering are proposed in Törn (1983).

Stochastic global optimization methods, like genetic algorithms, can also be applied in multiobjective optimization. An example of this approach is given in Osyczka and Kundu (1995). Another possibility for avoiding jamming into locally Pareto optimal solutions is to use simulated annealing or tabu search as an underlying solver.

In Arbel and Korhonen (1996a, 1997a), a new aspiration level-based method is developed in the spirit of interior point methods (of linear programming) for MOLP problems. The idea is to wander in the interior of the feasible objective region and only at the end to ascend to the Pareto optimal surface. Here, the generally adopted idea that decision makers should handle only (weakly) Pareto optimal solutions is called into question. One can justify such an approach by the fact that the decision maker can see some improvement in each objective function instead of having to trade off all the time. The interior point method used is an affine-scaling primal algorithm (also treated in Arbel (1993, 1994b, c)). The same idea is implemented by using an interior point method called the primal-dual algorithm in Arbel and Korhonen (1996b, 1997b) (also treated in Arbel (1994a, 1995)). Another modification of interior point methods for MOLP problems is described in Arbel and Oren (1994, 1996). In this method, the gradient of an implicitly known value function is approximated and a method of multiattribute decision analysis (namely AHP) is employed in comparing alternatives. The gradients of an implicitly known value function are also approximated and primal-dual linear methods used in Arbel (1997).

Results from other fields of research, for example, game theory, can also be used in the solution processes. Among others, Rao studies the relationship between Pareto optimal solutions and game theory in Rao (1987). He also applies his results to structural optimization.

Another important area of development is software designed to implement different methods and, especially, the user interface. As has been demonstrated, few well-known software products exist for nonlinear multiobjective optimization problems. As more and more advanced computers and graphical devices are created, more tools become available in the quest for ease and even enjoyment of use. This in turn involves new ideas for representing information, such as illustrating alternatives in natural language phrases or using new kinds of symbols. If the interface is able to adapt to the decision maker's style of making decisions and is of help in analyzing the alternatives and results, and can perhaps give suggestions or advice, then the interface may even overcome some of the deficiencies of the method itself.

As far as large-scale problems are concerned, the possibilities of parallel computing are worth examining in making the solution processes more efficient. Multimedia possibilities in decision support systems are reviewed, for example, in Grauer and Merten (1995).

One potentiality not to be forgotten is the utilization of expert systems. These can be applied to both suggesting a solution method according to the properties of the problem and the preferences of the decision maker, and in the solution process itself through supporting the decision maker. As an example, interactive MOLP methods and expert system techniques are integrated in Antunes et al. (1992b). The system described includes five methods, among them, STEM and the ZW method. When the user of the system expresses her or his hopes for further actions (such as a wish to get to know the neighbourhood of the current solution), the system suggests one of the available interactive methods. Computer graphics are also available. There are many features that deserve further research and development, but this is certainly an interesting path to follow.

A way of utilizing artificial neural networks in developing interactive multiobjective optimization methods in proposed in Sun et al. (1996). The decision maker is asked to articulate preference information over representative samples of the Pareto optimal set and the neural network is trained to represent this preference structure. The neural network is then used to generate improved solutions. The preference information can be specified by a preference value for each alternative or by pairwise comparisons between the alternatives. Possibilities of artificial intelligence and neural networks in multiobjective optimization are also charted in Gal and Hanne (1997).

The possibilities of several new technologies in computer science are reviewed in Antunes and Tsoukiàs (1997). The topics handled are fuzzy sets, multimedia, distributed computing, expert systems, object-oriented programming, neural networks, and the World-Wide Web. For example, the possibilities of the World-Wide Web in implementing interactive methods and making them easily available were dealt with in Subsection 5.12.8 of Part II. The example given was WWW-NIMBUS.

One more thing to mention are spreadsheets. They are widely used and thus provide a familiar environment for implementing interactivity in the methods. This idea is realized in Steuer (1997) but it deserves further examination.

Flexibility in the mathematical modelling of the problem is often desirable. Flexibility includes the possibility of interchanging the roles of objective and constraint functions and updating the model if necessary. The decision maker may, for example, wish to relax some constraints in order to be able to attain certain aspirations. This means that integrating the modelling and the solution processes deserves more attention.

It is not to be forgotten that dealing with incomplete information or uncertainty is a part of solving real-life applications. This area is important even though it has not been included in this book.

5. EPILOGUE

We have presented a self-contained survey of the state of the art of nonlinear multiobjective optimization together with a great number of further references. After treating several important concepts and their relations, we have considered some theoretical results and connections.

We have demonstrated the methodology of multiobjective optimization by describing several methods and by giving references in respect of a large number of other methods. Methods have been classified into four groups according to the contribution of the decision maker in the solution process. Because the group of interactive methods has been developed most, it has received the main emphasis. We have endeavoured to characterize the methods by some comments on their positive and negative features.

Some of the features of the interactive methods dealt with have been collected in a comparative table. Selected experiences and comparative observations of the methods have also been presented. In addition, some attempts to aid in the selection of a solution method have been made. A decision tree containing interactive methods has been suggested.

Some software packages have been mentioned. As far as software is concerned, several possibilities of graphical illustrations of alternative solutions have been introduced.

In general, one can say that the theory and the methods of multiobjective optimization have been extensively developed during the past couple of decades. Software implementations are considerably less in evidence. There is also a lack of documentation in solving real-life multiobjective optimization problems (using more developed methods). The reasons for this may be ignorance of the full range of possibilities contained in existing methods as well as the lack of suitable methods. For our part, we have filled a gap in the literature by collecting several nonlinear multiobjective optimization methods between the same covers.

In the development of methods the obvious conclusion is that it is important to continue in the direction of user-friendliness. Methods must be even better able to correspond to the characteristics of the decision maker. If the aspirations of the decision maker change during the solution process, the algorithm must be able to cope with this situation. Computational tests have confirmed the idea that decision makers want to feel in control of the solution process, and

K. Miettinen, *Nonlinear Multiobjective Optimization*

consequently they must understand what is happening. However, sometimes the decision maker simply needs support, and this should be available as well. Thus, the aim is to have methods that support learning so that guidance is given whenever necessary. The decision maker must be the basis in developing new interactive methods. Specific methods for different areas of application that take into account the characteristics of the problems are also important.

REFERENCES

Agrell P.J., *On Redundancy in Multi Criteria Decision Making*, European Journal of Operational Research **98**, No. 3 (1997), 571–586.

Agrell P.J., Lence B.J., Stam A., *An Interactive Multicriteria Decision Model for Multipurpose Reservoir Management: the Shellmouth Reservoir*, Journal of Multi-Criteria Decision Analysis **7**, No. 2 (1998), 61–86.

Aksoy Y., *Interactive Multiple Objective Decision Making: A Bibliography (1965–1988)*, Management Research News **13**, No. 2 (1990), 1–8.

Aksoy Y., Butler T.W., Minor E.D. III, *Comparative Studies in Interactive Multiple Objective Mathematical Programming*, European Journal of Operational Research **89**, No. 2 (1996), 408–422.

Al-alvani J.E., Hobbs B.F., Malakooti B., *An Interactive Integrated Multiobjective Optimization Approach for Quasiconcave/Quasiconvex Utility Functions*, Multiple Criteria Decision Making: Proceedings of the Ninth International Conference: Theory and Applications in Business, Industry, and Government, Edited by A. Goicoechea, L. Duckstein, S. Zionts, Springer-Verlag, New York, 1992, pp. 45–60.

Alekseichik M.I., Naumov G.E., *Existence of Efficient Points in Infinite-Dimensional Vector Optimization Problem*, Automation and Remote Control **42**, No. 5, Part 2 (1981), 666–670.

Angehrn A.A., *Supporting Multicriteria Decision Making: New Perspectives & New Systems*, Working Paper, INSEAD, Fontainebleau (1990a).

Angehrn A.A., *"Triple C" A Visual Interactive MCDSS*, Working Paper, INSEAD, Fontainebleau (1990b).

Antunes C.H., Tsoukiàs A., *Against Fashion: A Travel Survival Kit in "Modern" MCDA*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 378–389.

Antunes C.H., Alves M.J., Silva A.L., Clímaco J.N., *An Integrated MOLP Method Base Package - A Guided Tour of TOMMIX*, Computers & Operations Research **19**, No. 7 (1992a), 609–625.

Antunes C.H., Melo M.P., Clímaco J.N., *On the Integration of an Interactive MOLP Procedure Base and Expert System Technique*, European Journal of Operational Research **61**, No. 1–2 (1992b), 135–144.

Arbel A., *Generating Interior Search Directions for Multiobjective Linear Programming Using Approximate Gradients and Efficient Anchoring Points*, Optimization **28**, No. 2 (1993), 149–164.

Arbel A., *An Interior Multiobjective Primal-Dual Linear Programming Algorithm Using Approximated Gradients and Sequential Generation of Anchor Points*, Optimization **30**, No. 2 (1994a), 137–150.

Arbel A., *Anchoring Points and Cones of Opportunities in Interior Multiobjective Linear Programming*, Journal of the Operational Research Society **45**, No. 1 (1994b), 83–96.

Arbel A., *Using Efficient Anchoring Points for Generating Search Directions in Interior Multiobjective Linear Programming*, Journal of the Operational Research Society **45**, No. 3 (1994c), 330–344.

Arbel A., *An Interior Multiple Objective Primal-Dual Linear Programming Algorithm Using Efficient Anchoring Points*, Journal of the Operational Research Society **46**, No. 9 (1995), 1121–1132.

Arbel A., *An Interior Multiobjective Primal-Dual Linear Programming Algorithm Based on Approximated Gradients and Efficient Anchoring Points*, Computers & Operations Research **24**, No. 4 (1997), 353–365.

Arbel A., Korhonen P., *Using Aspiration Levels in an Interactive Interior Multiobjective Linear Programming Algorithm*, European Journal of Operational Research **89**, No. 1 (1996a), 193–201.

Arbel A., Korhonen P., *Using Aspiration Levels in an Interior Primal-Dual Multiobjective Linear Programming Algorithm*, Journal of Multi-Criteria Decision Analysis **5**, No. 1 (1996b), 61–71.

Arbel A., Korhonen P., *Generating Interior Search Directions in Multiple Objective Linear Programming Problems Using Aspiration Levels*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997a, pp. 146–155.

Arbel A., Korhonen P., *An Interior Multiobjective Linear Programming Algorithm Using Aspirations*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997b, pp. 245–254.

Arbel A., Oren S.S., *A Modification of Karmarkar's Algorithm to Multiple Objective Linear Programming Problems*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 37–46.

Arbel A., Oren S.S., *Using Approximate Gradients in Developing an Interactive Interior Primal-Dual Multiobjective Linear Programming Algorithm*, European Journal of Operational Research **89**, No. 1 (1996), 202–211.

Armann R., *Solving Multiobjective Programming Problems by Discrete Representation*, Optimization **20**, No. 4 (1989), 483–492.

Athan T.W., Papalambros P.Y., *Multicriteria Optimization of ABS Control Algorithms Using a Quasi-Monte Carlo Method*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 457–466.

Aubin J.-P., Näslund B., *An Exterior Branching Algorithm*, Working Paper 72-42, European Institute for Advanced Studies in Management, Brussels (1972).

Baba N., Takeda H., Miyake T., *Interactive Multi-Objective Programming Technique Using Random Optimization Method*, International Journal of Systems Science **19**, No. 1 (1988), 151–159.

Balachandran M., Gero J.S., *The Noninferior Set Estimation (NISE) Method for Three Objective Problem*, Engineering Optimization **9**, No. 2 (1985), 77–88.

Balbás A., Guerra P.J., *Sensitivity Analysis for Convex Multiobjective Programming in Abstract Spaces*, Journal of Mathematical Analysis and Applications **202**, No. 2 (1996), 645–658.

Ballestero E., *Utility Functions: A Compromise Programming Approach to Specification and Optimization*, Journal of Multi-Criteria Decision Analysis **6**, No. 1 (1997a), 11–16.

Ballestero E., *Selecting the CP Metric: A Risk Aversion Approach*, European Journal of Operational Research **97**, No. 3 (1997b), 593–596.

Ballestero E., Romero C., *A Theorem Connecting Utility Function Optimization and Compromise Programming*, Operations Research Letters **10**, No. 7 (1991), 421–427.

Bard J.F., *A Multiobjective Methodology for Selecting Subsystem Automation Options*, Management Science **32**, No. 12 (1986), 1628–1641.

Bárdossy A., Bogárdi I., Duckstein L., *Composite Programming as an Extension of Compromise Programming*, Mathematics of Multi Objective Optimization, Edited by P. Serafini, Springer-Verlag, Wien, New York, 1985, pp. 375–408.

Batishchev D.I., Anuchin V.F., Shaposhnikov D.E., *The Use of the Qualitative Information on the Importance of Particular Criteria for the Computation of Weighting Coefficients*, Multiobjective Problems of Mathematical Programming, Edited by A. Lewandowski, V. Volkovich, Lecture Notes in Economics and Mathematical Systems 351, Springer-Verlag, 1991, pp. 2–7.

Bazaraa M.S., Sherali H.D., Shetty C.M., *Nonlinear Programming: Theory and Algorithms*, Second Edition, John Wiley & Sons, Inc., New York, 1993.

Belkeziz K., Pirlot M., *Proper Efficiency in Nonconvex Vector-Maximization-Problems*, European Journal of Operational Research **54**, No. 1 (1991), 74–80.

Bell D.E., *Double-Exponential Utility Functions*, Mathematics of Operations Research **11**, No. 2 (1986), 351–361.

Bell D.E., *Risk, Return, and Utility*, Management Science **41**, No. 1 (1995), 23–30.

Bell P.C., O'Keefe R.M., *An Experimental Investigation into the Efficacy of Visual Interactive Simulation*, Management Science **41**, No. 6 (1995), 1018–1038.

Belton V., Vickers S., *Use of a Simple Multi-Attribute Value Function Incorporating Visual Interactive Sensitivity Analysis for Multiple Criteria Decision Making*, Readings in Multiple Criteria Decision Aid, Edited by C.A. Bana e Costa, Springer-Verlag, Berlin, Heidelberg, 1990, pp. 319–334.

Benayoun R., de Montgolfier J., Tergny J., Laritchev O., *Linear Programming with Multiple Objective Functions: Step Method (STEM)*, Mathematical Programming **1**, No. 3 (1971), 366–375.

Benbasat I., Dexter A.S., *An Experimental Evaluation of Graphical and Color-Enhanced Information Presentation*, Management Science **31**, No. 11 (1985), 1348–1364.

Benito-Alonso M.A., Devaux P., *Location and Size of Day Nurseries – a Multiple Goal Approach*, European Journal of Operational Research **6**, No. 2 (1981), 195–198.

Benoist J., *Connectedness of the Efficient Set for Strictly Quasiconcave Sets*, Journal of Optimization Theory and Applications **96**, No. 3 (1998), 627–654.

Benson H.P., *Existence of Efficient Solutions for Vector Maximization Problems*, Journal of Optimization Theory and Applications **26**, No. 4 (1978), 569–580.

Benson H.P., *An Improved Definition of Proper Efficiency for Vector Maximization with Respect to Cones*, Journal of Mathematical Analysis and Applications **71**, No. 1 (1979a), 232–241.

Benson H.P., *Vector Maximization with Two Objective Functions*, Journal of Optimization Theory and Applications **28**, No. 2 (1979b), 253–257.

Benson H.P., *On a Domination Property for Vector Maximization with Respect to Cones*, Journal of Optimization Theory and Applications **39**, No. 1 (1983), 125–132; *Errata Corrige*, Journal of Optimization Theory and Applications **43**, No. 3 (1984), 477–479.

Benson H.P., *Complete Efficiency and the Initialization of Algorithms for Multiple Objective Programming*, Operations Research Letters **10**, No. 8 (1991), 481–487.

Benson H.P., Aksoy Y., *Using Efficient Feasible Directions in Interactive Multiple Objective Linear Programming*, Operations Research Letters **10**, No. 4 (1991), 203–209.

Benson H.P., Morin T.L., *The Vector Maximization Problem: Proper Efficiency and Stability*, SIAM Journal on Applied Mathematics **32**, No. 1 (1977), 64–72.

Benson H.P., Sayin S., *Optimization over the Efficient Set: Four Special Cases*, Journal of Optimization Theory and Applications **80**, No. 1 (1994), 3–18.

Benson H.P., Sayin S., *Towards Finding Global Representations of the Efficient Set in Multiple Objective Mathematical Programming*, Naval Research Logistics **44**, No. 1 (1997), 47–67.

Ben-Tal A., *Characterization of Pareto and Lexicographic Optimal Solutions*, Multiple Criteria Decision Making Theory and Application, Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 177, Springer-Verlag, Berlin, Heidelberg, 1980, pp. 1–11.

Berbel J., Zamora R., *Application of MOP and GP to Wildlife Management (Deer). A Case in the Mediterranean Ecosystem*, Multi-Objective Programming and Goal Programming: Theories and Applications, Edited by M. Tamiz, Lecture Notes in Economics and Mathematical Systems 432, Springer-Verlag, Berlin, Heidelberg, 1996, pp. 300–308.

Bestle D., Eberhard P., *Dynamic System Design via Multicriteria Optimization*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 467–478.

Bhatia D., Aggarwal S., *Optimality and Duality for Multiobjective Nonsmooth Programming*, European Journal of Operational Research **57**, No. 3 (1992), 360–367.

Bhatia D., Datta N., *Necessary Conditions and Subgradient Duality for Non-Differentiable and Non-Convex Multi-Objective Programming Problem*, Cahiers du Centre d'Études de Recherche Opérationelle **27**, No. 1–2 (1985), 131–139.

Bischoff E., *Two Empirical Tests with Approaches to Multiple-Criteria Decision Making*, Plural Rationality and Interactive Decision Processes, Edited by M. Grauer, M. Thompson, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 248, Springer-Verlag, Berlin, Heidelberg, 1985, pp. 344–347.

Bischoff E., *Multi-Objective Decision Analysis – The Right Objectives?*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986, pp. 155–160.

Bitran G.R., *Duality for Nonlinear Multiple-Criteria Optimization Problems*, Journal of Optimization Theory and Applications **35**, No.3 (1981), 367–401.

Bogetoft P., Hallefjord Å., Kok M., *On the Convergence of Reference Point Methods in Multiobjective Programming*, European Journal of Operational Research **34**, No. 1 (1988), 56–68.

Borwein J.M., *Proper Efficient Points for Maximizations with Respect to Cones*, SIAM Journal on Control and Application **15**, No. 1 (1977), 57–63.

Borwein J.M., *On the Existence of Pareto Efficient Points*, Mathematics of Operations Research **8**, No. 1 (1983), 64–73.

Borwein J.M., Zhuang D.M., *Super Efficiency in Convex Vector Optimization*, ZOR – Methods and Models of Operations Research **35**, No. 3 (1991), 175–184.

Borwein J.M., Zhuang D.M., *Super Efficiency in Vector Optimization*, Transactions of the American Mathematical Society **338**, No. 1 (1993), 105–122.

Bowman V.J. Jr., *On the Relationship of the Tchebycheff Norm and the Efficient Frontier of Multiple-Criteria Objectives*, Multiple Criteria Decision Making, Edited by H. Thiriez, S. Zionts, Lecture Notes in Economics and Mathematical Systems 130, Springer-Verlag, Berlin, Heidelberg, 1976, pp. 76–85.

Brans J.-P., *The Space of Freedom of the Decision Maker – Modelling the Human Brain*, European Journal of Operational Research **92**, No. 3 (1996), 593–602.

Brans J.P., Mareschal B., *The PROMETHEE Methods for MCDM, the PROMCALC, GAIA and BANKADVISER Software*, Readings in Multiple Criteria Decision Aid, Edited by C.A. Bana e Costa, Springer-Verlag, Berlin, Heidelberg, 1990, pp. 216–252.

Brauer D.C., Naadimuthu G., *A Goal Programming Model for Aggregate Inventory and Distribution Planning*, Mathematical and Computer Modelling **16**, No. 3 (1992), 81–90.

Brockhoff K., *Experimental Test of MCDM Algorithms in a Modular Approach*, European Journal of Operational Research **22**, No. 2 (1985), 159–166.

Brosowski B., da Silva A.R., *Simple Tests for Multi-Criteria Optimality*, OR Spektrum **16**, No. 4 (1994), 243–247.

Buchanan J.T., *Multiple Objective Mathematical Programming: A Review*, New Zealand Operational Research **14**, No. 1 (1986), 1–27.

Buchanan J.T., *An Experimental Evaluation of Interactive MCDM Methods and the Decision Making Process*, Journal of the Operational Research Society **45**, No. 9 (1994), 1050–1059.

Buchanan J.T., *A Naïve Approach for Solving MCDM Problems: The GUESS Method*, Journal of the Operational Research Society **48**, No. 2 (1997), 202–206.

Buchanan J.T., Corner J.L., *The Effects of Anchoring in Interactive MCDM Solution Methods*, Computers & Operations Research **24**, No. 10 (1997), 907–918.

Buchanan J.T., Daellenbach H.G., *A Comparative Evaluation of Interactive Solution Methods for Multiple Objective Decision Models*, European Journal of Operational Research **29**, No. 3 (1987), 353–359.

Buede D., *Second Overview of the MCDA Software Market*, Journal of Multi-Criteria Decision Analysis **5**, No. 4 (1996), 312–316.

Bushenkov V., Kaitala V., Lotov A., Pohjola M., *Decision and Negotiation Support for Transboundary Air Pollution Control between Finland, Russia and Estonia*, Finnish Economic Papers **7**, No. 1 (1994).

Bushenkov V.A., Chernykh O.L., Kamenev G.K., Lotov A.V., *Multidimensional Images Given by Mappings: Construction and Visualization*, Pattern Recognition and Image Analysis **5**, No. 1 (1995), 35–56.

Caballero R., Rey L., Ruiz F., *Determination of Satisfying and Efficient Solutions in Convex Multi-Objective Programming*, Optimization **37**, No. 2 (1996), 125–137.

Caballero R., Rey L., Ruiz F., González M., *An Algorithmic Package for the Resolution and Analysis of Convex Multiple Objective Problems*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 275–284.

Cambini A., Martein L., *On the Existence of Efficient Points*, Optimization **28**, No. 3-4 (1994), 283–290.

Carmichael D.G., *Structural Modelling and Optimization*, Ellis Horwood Ltd., 1981.

Censor Y., *Pareto Optimality in Multiobjective Problems*, Applied Mathematics and Optimization **4**, No. 1 (1977), 41–59.

Chankong V., Haimes Y.Y., *The Interactive Surrogate Worth Trade-Off (ISWT) Method for Multiobjective Decision-Making*, Multiple Criteria Problem Solving, Edited by S. Zionts, Lecture Notes in Economics and Mathematical Systems 155, Springer-Verlag, Berlin, Heidelberg, 1978, pp. 42–67.

Chankong V., Haimes Y.Y., *On the Characterization of Noninferior Solutions of the Vector Optimization Problem*, Automatica **18**, No. 6 (1982), 697–707.

Chankong V., Haimes Y.Y., *Optimization-Based Methods for Multiobjective Decision-Making: An Overview*, Large Scale Systems **5**, No. 1 (1983a), 1–33.

Chankong V., Haimes Y.Y., *Multiobjective Decision Making Theory and Methodology*, Elsevier Science Publishing Co., Inc., New York, 1983b.

Chankong V., Haimes Y.Y., Thadathil J., Zionts S., *Multiple Criteria Optimization: A State of the Art Review*, Decision Making with Multiple Objectives, Lecture Notes in Economics and Mathematical Systems 242, Edited by Y.Y. Haimes, V. Chankong, Springer-Verlag, 1985, pp. 36–90.

Charnes A., Cooper W.W., *Management Models and Industrial Applications of Linear Programming*, vol. 1, John Wiley & Sons, Inc., New York, 1961.

Charnes A., Cooper W.W., *Goal Programming and Multiple Objective Optimization; Part 1*, European Journal of Operational Research **1**, No. 1 (1977), 39–54.

Charnes A., Cooper W.W., Ferguson R.O., *Optimal Estimation of Executive Compensation by Linear Programming*, Management Science **1**, No. 2 (1955), 138–151.

Chen G.-Y., *Necessary Conditions of Nondominated Solutions in Multicriteria Decision Making*, Journal of Mathematical Analysis and Applications **104**, No. 1 (1984), 38–46.

Chen T., Wang B.-I., *An Interactive Method for Multiobjective Decision Making*, Large Scale Systems: Theory and Applications 1983, Edited by A. Straszak, Pergamon Press, 1984, pp. 277–282.

Chew K.L., Choo E.U., *Pseudolinearity and Efficiency*, Mathematical Programming **28**, No. 2 (1984), 226–239.

Choo E.U., Atkins D.R., *Proper Efficiency in Nonconvex Multicriteria Programming*, Mathematics of Operations Research **8**, No. 3 (1983), 467–470.

Chou J.H., Hsia W.-S., Lee T.-Y., *On Multiple Objective Programming Problems with Set Functions*, Journal of Mathematical Analysis and Applications **105**, No. 2 (1985), 383–394.

Clarke F.H., *Optimization and Nonsmooth Analysis*, John Wiley & Sons, Inc., 1983.

Cleveland W.S., *The Elements of Graphing Data*, Revised Edition, AT&T Bell Laboratories, Murray Hill, New Jersey, 1994.

Clímaco J.N., Antunes C.H., *Flexible Method Bases and Man-Machine Interfaces as Key Features in Interactive MOLP Approaches*, Multiple Criteria Decision Support, Edited by P. Korhonen, A. Lewandowski, J. Wallenius, Lecture Notes in Economics and Mathematical Systems 356, Springer-Verlag, 1991, pp. 207–216.

Clímaco J.N., Antunes C.H., *Man-Machine Interfacing in MCDA*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 239–253.

Clímaco J.N., Antunes C.H., Alves M.J., *From TRIMAP to SOMMIX – Building Effective Interactive MOLP Computational Tools*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 285–296.

Clinton R.J., Troutt M.D., *Toward Optimal Scaling in the Method of Abstract Forces for Interactive Multiple Criteria Optimization*, Mathematical and Computer Modelling **11** (1988), 589–594.

Cohon J.L., *Multiobjective Programming and Planning*, Academic Press, Inc., New York, 1978.

Cohon J.L., *Multicriteria Programming: Brief Review and Application*, Design Optimization, Edited by J.S. Gero, Academic Press, Inc., 1985, pp. 163–191.

Colson G., de Bruyn C., *Multiple Criteria Supports on Micro-Computers*, Belgian Journal of Operations Research, Statistics and Computer Science **27**, No. 1 (1987), 29–58.

Corley H.W., *A New Scalar Equivalence for Pareto Optimization*, IEEE Transactions on Automatic Control **25**, No. 4 (1980), 829–830.

Corner J.L., Buchanan J.T., *Experimental Consideration of Preference in Decision Making under Certainty*, Journal of Multi-Criteria Decision Analysis **4**, No. 2 (1995), 107–121.

Corner J.L., Buchanan J.T., *Capturing Decision Maker Preference: Experimental Comparison of Decision Analysis and MCDM Techniques*, European Journal of Operational Research **98**, No. 1 (1997), 85–97.

Corner J.L., Corner P.D., *Characteristics of Decisions in Decision Analysis Practice*, Journal of the Operational Research Society **46**, No. 3 (1995), 304–314.

Costa J.P., Clímaco J.N., *A Multiple Reference Point Parallel Approach in MCDM*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 255–263.

Crama Y., *Analysis of STEM-Like Solutions to Multi-Objective Programming Problems*, Multi-Objective Decision Making, Edited by S. French, R. Hartley, L.C. Thomas, D.J. White, Academic Press, 1983, pp. 208–213.

Craven B.D., *On Sensitivity Analysis for Multicriteria Optimization*, Optimization **19**, No. 4 (1988), 513–523.

Craven B.D., *Nonsmooth Multiobjective Programming*, Numerical Functional Analysis and Optimization **10**, No. 1–2 (1989), 49–64.

Current J.R., Storbeck J.E., *A Multiobjective Approach to Design Franchise Outlet Networks*, Journal of the Operational Research Society **45**, No. 1 (1994), 71–81.

Da Cunha N.O., Polak E., *Constrained Minimization under Vector-Valued Criteria in Finite Dimensional Spaces*, Journal of Mathematical Analysis and Applications **19**, No. 1 (1967), 103–124.

Das I., Dennis J.E., *A Closer Look at Drawbacks of Minimizing Weighted Sums of Objectives for Pareto Set Generation in Multicriteria Optimization Problems*, Structural Optimization **14**, No. 1 (1997), 63–69.

Das I., Dennis J.E., *Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems*, SIAM Journal on Optimization **8**, No. 3 (1998), 631–657.

Das L.N., Nanda S., *Symmetric Dual Multiobjective Programming*, European Journal of Operational Research **97**, No. 1 (1997), 167–171.

Dauer J.P., Fosnaugh T.A., *Optimization over the Efficient Set*, Journal of Global Optimization **7**, No. 3 (1995), 261–277.

Dauer J., Liu Y.-H., *Multi-Criteria and Goal Programming*, Advances in Sensitivity Analysis and Parametric Programming, Edited by T. Gal. H.J. Greenberg, Kluwer Academic Publishers, Boston, 1997, pp. 11-1–11-31.

Dauer J.P., Osman M.S.A., *Decomposition of the Parametric Space in Multiobjective Convex Programs Using the Generalized Tchebycheff Norm*, Journal of Mathematical Analysis and Applications **107**, No. 1 (1985), 156–166.

Dauer J.P., Stadler W., *A Survey of Vector Optimization in Infinite-Dimensional Spaces, Part 2*, Journal of Optimization Theory and Applications **51**, No. 2 (1986), 205–241.

de Kluyver C.A., *An Exploration of Various Goal Programming Formulations – with Application to Advertising Media Selection*, Journal of the Operational Research Society **30**, No. 2 (1979), 167–171.

de Neufville R., McCord M., *Unreliable Measurement of Utility: Significant Problems for Decision Analysis*, Operational Research '84, Edited by J.P. Brans, Elsevier Science Publishers B.V. (North-Holland), Amsterdam, 1984, pp. 464–476.

Deng S., *Characterizations of the Nonemptiness and Compactness of Solution Sets in Convex Vector Optimization*, Journal of Optimization Theory and Applications **96**, No. 1 (1998a), 123–131.

Deng S., *On Efficient Solutions in Vector Optimization*, Journal of Optimization Theory and Applications **96**, No. 1 (1998b), 201–209.

Despontin M., Moscarola J., Spronk J., *A User-Oriented Listing of Multiple Criteria Decision Methods*, Revue Belge de Statistique, d'Informatique et de Recherche Operationnelle **23**, No. 4 (1983), 4–110.

Dessouky M.I., Ghiassi M., Davis W.J., *Estimates of the Minimum Nondominated Criterion Values in Multiple-Criteria Decision-Making*, Engineering Costs and Production Economics **10** (1986), 95–104.

Diaz A., *Interactive Solution to Multiobjective Optimization Problems*, International Journal for Numerical Methods in Engineering **24**, No. 10 (1987), 1865–1877.

Dinkelbach W., *Multicriteria Decision Models with Specified Goal Levels*, Multiple Criteria Decision Making Theory and Application, Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 177, Springer-Verlag, Berlin, Heidelberg, 1980, pp. 52–59.

Doležal J., *Necessary Conditions for Pareto Optimality in Nondifferentiable Problems*, Problems of Control and Information Theory **14**, No. 2 (1985), 131–140.

Durier R., *Weighting Factor Results in Vector Optimization*, Journal of Optimization Theory and Applications **58**, No. 3 (1988), 411–430.

Dyer J.S., *Interactive Goal Programming*, Management Science **19**, No. 1 (1972), 62–70.

Dyer J.S., *A Time-Sharing Computer Program for the Solution of the Multiple Criteria Problem*, Management Science **19**, No. 12 (1973a), 1379–1383.

Dyer J.S., *An Empirical Investigation of a Man-Machine Interactive Approach to the Solution of the Multiple Criteria Problem*, Multiple Criteria Decision Making, Edited by J.L. Cochrane, M. Zeleny, University of South Carolina Press, Columbia, South Carolina, 1973b, pp. 202–216.

Dyer J.S., *The Effects of Errors in the Estimation of the Gradient on the Frank-Wolfe Algorithm, with Implications for Interactive Programming*, Operations Research **22**, No. 1 (1974), 160–174.

Dyer J.S., Sarin R.K., *Multicriteria Decision Making*, Mathematical Programming for Operations Researchers and Computer Scientists, Edited by A.G. Holzman, Marcel Dekker, Inc., 1981, pp. 123–148.

Dyer J.S., Fishburn P.C., Steuer R.E., Wallenius J., Zionts S., *Multiple Criteria Decision Making, Multiattribute Utility Theory: The Next Ten Years*, Management Science **38**, No. 5 (1992), 645–654.

Ecker J.G., Kouada I.A., *Finding Efficient Points for Linear Multiple Objective Programs*, Mathematical Programming **8**, No. 3 (1975), 375–377.

Edgeworth F.Y., *Mathematical Psychics*, University Microfilms International (Out-of-Print Books on Demand), 1987 (the original edition in 1881).

Egudo R.R., *A Duality Theorem for a Fractional Multiobjective Problem with Square Root Terms*, Recent Developments in Mathematical Programming, Edited by S. Kumar, Gordon and Breach Science Publishers, 1991, pp. 101–113.

Ehrgott M., Klamroth K., *Connectedness of Efficient Solutions in Multiple Criteria Combinatorial Optimization*, European Journal of Operational Research **97**, No. 1 (1997), 159–166.

Eiduks J., *Control of Complicated Systems in Case of Vectorial Indicator of Quality*, Fourth Formator Symposium on Mathematical Methods for the Analysis of Large-Scale Systems, Edited by J. Bebeš, L. Bakule, Academia, Prague, 1983, pp. 165–180.

Eiselt H.A., Pederzoli G., Sandblom C.-L., *Continuous Optimization Models*, Operations Research: Theory, Techniques, Applications, Walter de Gruyter & Co., Berlin, New York, 1987.

El Abdouni B., Thibault L., *Lagrange Multipliers for Pareto Nonsmooth Programming Problems in Banach Spaces*, Optimization **26**, No. 3-4 (1992), 277–285.

Eom H.B., *The Current State of Multiple Criteria Decision Support Systems*, Human Systems Management **8** (1989), 113–119.

Eschenauer H.A., *Multicriteria Optimization Procedures in Application on Structural Mechanics Systems*, Recent Advances and Historical Developments of Vector Optimization, Edited by J. Jahn, W. Krabs, Lecture Notes in Economics and Mathematical Systems 294, Springer-Verlag, Berlin, Heidelberg, 1987, pp. 345–376.

Eschenauer H.A., Schäfer E., Bernau H., *Application of Interactive Vector Optimization Methods with Regard to Problems in Structural Mechanics*, Discretization Methods and Structural Optimization – Procedures and Applications, Edited by H.A. Eschenauer, G. Thierauf, Lecture Notes in Engineering 42, Springer-Verlag, Berlin, Heidelberg, 1989, pp. 95–101.

Eschenauer H., Koski J., Osyczka A. (Eds.), *Multicriteria Design Optimization Procedures and Applications*, Springer-Verlag, Berlin, Heidelberg, 1990a.

Eschenauer H.A., Osyczka A., Schäfer E., *Interactive Multicriteria Optimization in Design Process*, Multicriteria Design Optimization Procedures and Applications, Edited by H. Eschenauer, J. Koski, A. Osyczka, Springer-Verlag, Berlin, Heidelberg, 1990b, pp. 71–114.

Ester J., *About the Stability of the Solution of Multiple Objective Decision Making Problems with Generalized Dominating Sets*, Systems Analysis Modelling Simulation **1**, No. 4 (1984), 287–291.

Ester J., Holzmüller R., *About Some Methods and Applications in Multicriteria Decision Making*, Systems Analysis Modelling Simulation **3**, No. 5 (1986), 425–438.

Ester J., Tröltzsch F., *On Generalized Notions of Efficiency in Multicriteria Decision Making*, Systems Analysis Modelling Simulation **3**, No. 2 (1986), 153–161.

Evans G.W., *An Overview of Techniques for Solving Multiobjective Mathematical Programs*, Management Science **30**, No. 11 (1984), 1268–1282.

Federer H., *Geometric Measure Theory*, Springer-Verlag, Berlin, Heidelberg, 1969.

Ferreira P.A.V., Geromel J.C., *An Interactive Projection Method for Multicriteria Optimization Problems*, IEEE Transactions on Systems, Man, and Cybernetics **20**, No. 3 (1990), 596–605.

Ferreira P.A.V., Machado M.E.S., *Solving Multiple-Objective Problems in the Objective Space*, Journal of Optimization Theory and Applications **89**, No. 3 (1996), 659–680.

Fishburn P.C., *Lexicographic Orders, Utilities and Decision Rules: A Survey*, Management Science **20**, No. 11 (1974), 1442–1471.

Flavell R.B., *A New Goal Programming Formulation*, OMEGA, The International Journal of Management Science **4**, No. 6 (1976), 731–732.

Frank M., Wolfe P., *An Algorithm for Quadratic Programming*, Naval Research Logistics Quarterly **3**, No. 1–2 (1956), 95–110.

French S., *Interactive Multi-Objective Programming: Its Aims, Applications and Demands*, Journal of the Operational Research Society **35**, No. 9 (1984), 827–834.

Friedman L., Mehrez A., *Towards a Theory of Determining Size in Multi-Objective Service Systems*, European Journal of Operational Research **63**, No. 3 (1992), 398–408.

Friesz T.L., *Multiobjective Optimization in Transportation: The Case of Equilibrium Network Design*, Organizations: Multiple Agents with Multiple Criteria, Edited by N.N. Morse, Lecture Notes in Economics and Mathematical Systems 190, Springer-Verlag, Berlin, Heidelberg, 1981, pp. 116–127.

Gal T., *On Efficient Sets in Vector Maximum Problems – A Brief Survey*, European Journal of Operational Research **24**, No. 2 (1986), 253–264.

Gal T., *Postoptimal Analyses, Parametric Programming, and Related Topics: Degeneracy, Multicriteria Decision Making, Redundancy*, Second Edition, Walter de Gruyter & Co., Berlin, 1995.

Gal T., Hanne T., *On the Development and Future Aspects of Vector Optimization and MCDM*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 130–145.

Gal T., Leberling H., *Redundant Objective Functions in Linear Vector Maximum Problems and Their Determination*, European Journal of Operational Research **1**, No. 3 (1977), 176–184.

Gal T., Wolf K., *Stability in Vector Maximization – A Survey*, European Journal of Operational Research **25**, No. 2 (1986), 169–182.

Gardiner L.R., Steuer R.E., *Unified Interactive Multiple Objective Programming*, European Journal of Operational Research **74**, No. 3 (1994a), 391–406.

Gardiner L.R., Steuer R.E., *Unified Interactive Multiple Programming: An Open Architecture for Accommodating New Procedures*, Journal of the Operational Research Society **45**, No. 12 (1994b), 1456–1466.

Gardiner L.R., Vanderpooten D., *Interactive Multiple Criteria Procedures: Some Reflections*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 290–301.

Gass S., Saaty T., *The Computational Algorithm for the Parametric Objective Function*, Naval Research Logistics Quarterly **2** (1955), 39–45.

Geoffrion A.M., *Proper Efficiency and the Theory of Vector Maximization*, Journal of Mathematical Analysis and Applications **22**, No. 3 (1968), 618–630.

Geoffrion A.M., Dyer J.S., Feinberg A., *An Interactive Approach for Multi-Criterion Optimization, with an Application to the Operation of an Academic Department*, Management Science **19**, No. 4 (1972), 357–368.

Geromel J.C., Ferreira P.A.V., *An Upper Bound on Properly Efficient Solutions on Multiobjective Optimization*, Operations Research Letters **10**, No. 2 (1991), 83–86.

Gershon M., Duckstein L., *An Algorithm for Choosing of a Multiobjective Technique*, Essays and Surveys on Multiple Criteria Decision Making, Edited by P. Hansen, Lecture Notes in Economics and Mathematical Systems 209, Springer-Verlag, Berlin, Heidelberg, 1983, pp. 53–62.

Ghosh D., Pal B.B., Basu M., *Implementation of Goal Programming in Long-Range Resource Planning in University Management*, Optimization **24**, No. 3–4 (1992), 373–383.

Giannikos I., El-Darzi E., Lees P., *An Integer Goal Programming Model to Allocate Offices to Staff in an Academic Institution*, Journal of the Operational Research Society **46**, No. 6 (1995), 713–720.

Gibson M., Bernardo J.J., Chung C., Badinelli R., *A Comparison of Interactive Multiple-Objective Decision Making Procedures*, Computers & Operations Research **14**, No. 2 (1987), 97–105.

Giokas D., Vassiloglou M., *A Goal Programming Model for Bank Assets and Liabilities Management*, European Journal of Operational Research **50**, No. 1 (1991), 48–60.

Goh C.J., Yang X.Q., *Analytic Efficient Solution Set for Multi-Criteria Quadratic Programs*, European Journal of Operational Research **92**, No. 1 (1996), 166–181.

Göpfert A., *Multicriterial Duality, Examples and Advances*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986, pp. 52–58.

Gouljashki V.G., Kirilov L.M., Narula S.C., Vassilev V.S., *A Reference Direction Interactive Algorithm of the Multiple Objective Nonlinear Integer Programming*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 308–317.

Granat J., Kreglewski K., Paczyński J., Stachurski A., *IAC-DIDASN++ Modular Modeling and Optimization System Theoretical Foundations*, Report of the Institute of Automatic Control, Warsaw University of Technology (1994a).

Granat J., Kreglewski K., Paczyński J., Stachurski A., *IAC-DIDASN++ Modular Modeling and Optimization System Users Guide*, Report of the Institute of Automatic Control, Warsaw University of Technology (1994b).

Grauer M., *A Dynamic Interactive Decision Analysis and Support System (DIDASS), User's Guide*, Working Paper WP-83-60, IIASA, Laxenburg (1983a).

Grauer M., *Reference Point Optimization – The Nonlinear Case*, Essays and Surveys on Multiple Criteria Decision Making, Edited by P. Hansen, Lecture Notes in Economics and Mathematical Systems 209, Springer-Verlag, Berlin, Heidelberg, 1983b, pp. 126–135.

Grauer M., Merten U., *Analysis of Multimedia Concepts for Decision Support*, Arbeitsberichte Nr. 20, Universität – Gesamthochschule – Siegen, 1995.

Grauer M., Lewandowski A., Wierzbicki A., *DIDASS – Theory, Implementation and Experiences*, Interactive Decision Analysis, Edited by M. Grauer, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 229, Springer-Verlag, 1984, pp. 22–30.

Graves S.B., Ringuest J.L., Bard J.F., *Recent Developments in Screening Methods for Nondominated Solutions in Multiobjective Optimization*, Computers & Operations Research **19**, No. 7 (1992), 683–694.

Guddat J., Guerra Vasquez F., Tammer K., Wendler K., *Multiobjective and Stochastic Optimization Based on Parametric Optimization*, Akademie-Verlag, Berlin, 1985.

Gulati T.R., Islam M.A., *Proper Efficiency in a Linear Fractional Vector Maximum Problem with Generalized Convex Constraints*, European Journal of Operational Research **36**, No. 3 (1988), 339–345.

Gulati T.R., Islam M.A., *Efficiency and Proper Efficiency in Nonlinear Vector Maximum Problems*, European Journal of Operational Research **44**, No. 3 (1990), 373–382.

Haimes Y.Y., *The Surrogate Worth Trade-off (SWT) Method and Its Extensions*, Multiple Criteria Decision Making Theory and Application, Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 177, Springer-Verlag, Berlin, Heidelberg, 1980, pp. 85–108.

Haimes Y.Y., *Multiple-Criteria Decisionmaking: A Retrospective Analysis*, IEEE Transactions on Systems, Man, and Cybernetics **15**, No. 3 (1985), 313–315.

Haimes Y.Y., Chankong V., *Kuhn-Tucker Multipliers as Trade-Offs in Multiobjective Decision-Making Analysis*, Automatica **15**, No. 1 (1979), 59–72.

Haimes Y.Y., Hall W.A., *Multiobjectives in Water Resource Systems Analysis: The Surrogate Worth Trade Off Method*, Water Resources Research **10**, No. 4 (1974), 615–624.

Haimes Y.Y., Li D., *Hierarchical Multiobjective Analysis for Large-Scale Systems: Review and Current Status*, Automatica **24**, No. 1 (1988), 53–69.

Haimes Y.Y., Lasdon L.S., Wismer D.A., *On a Bicriterion Formulation of the Problems of Integrated System Identification and System Optimization*, IEEE Transactions on Systems, Man, and Cybernetics **1** (1971), 296–297.

Haimes Y.Y., Hall W.A., Freedman H.T., *Multiobjective Optimization in Water Resources Systems*, Elsevier Scientific Publishing Company, Amsterdam, 1975.

Haimes Y.Y., Tarvainen K., Shima T., Thadathil J., *Hierarchical Multiobjective Analysis of Large-Scale Systems*, Hemisphere Publishing Corporation, New York, 1990.

Hall W.A., Haimes Y.Y., *The Surrogate Worth Trade-Off Method with Multiple Decision-Makers*, Multiple Criteria Decision Making Kyoto 1975, Edited by M. Zeleny, Lecture Notes in Economics and Mathematical Systems 123, Springer-Verlag, Berlin, Heidelberg, 1976, pp. 207–233.

Hallefjord Å., Jörnsten K., *An Entropy Target-Point Approach to Multiobjective Programming*, International Journal of Systems Science **17**, No. 4 (1986), 639–653.

Halme M., Korhonen P., *Nondominated Tradeoffs and Termination in Interactive Multiple Objective Linear Programming*, Improving Decision Making in Organisations, Edited by A.G. Lockett, G. Islei, Lecture Notes in Economics and Mathematical Systems 335, Springer-Verlag, Berlin, Heidelberg, 1989, pp. 410–423.

Harrison T.P., Rosenthal R.E., *Optimization of Utility and Value Functions*, Naval Research Logistics **35**, No. 3 (1988), 411–418.

Haslinger J., Neittaanmäki P., *Finite Element Approximation for Optimal Shape Design: Theory and Applications*, John Wiley & Sons, Inc., Chichester, 1988.

Hazen G.B., *Differential Characterizations of Nonconical Dominance in Multiple Objective Decision Making*, Mathematics of Operations Research **13**, No. 1 (1988), 174–189.

Helbig S., *On the Connectedness of the Set of Weakly Efficient Points of a Vector Optimization Problem in Locally Convex Spaces*, Journal of Optimization Theory and Applications **65**, No. 2 (1990), 257–270.

Helbig S., *Approximation of the Efficient Point Set by Perturbation of the Ordering Cone*, ZOR – Methods and Models of Operations Research **35**, No. 3 (1991), 197–220.

Hemaida R.S., Kwak N.K., *A Linear Goal Programming Model for Trans-Shipment Problems with Flexible Supply and Demand Constraints*, Journal of the Operational Research Society **45**, No. 2 (1994), 215–224.

Hemming T., *Multiobjective Decision Making under Certainty*, Doctoral Thesis, The Economic Research Institute at the Stockholm School of Economics, Stockholm, 1978.

Hemming T., *Some Modifications of a Large Step Gradient Method for Interactive Multicriterion Optimization*, Organizations: Multiple Agents with Multiple Criteria, Edited by N.N. Morse, Lecture Notes in Economics and Mathematical Systems 190, Springer-Verlag, Berlin, Heidelberg, 1981, pp. 128–139.

Henig M.I., *Existence and Characterization of Efficient Decisions with Respect to Cones*, Mathematical Programming **23**, No. 1 (1982a), 111–116.

Henig M.I., *Proper Efficiency with Respect to Cones*, Journal of Optimization Theory and Applications **36**, No. 3 (1982b), 387–407.

Henig M.I., *The Domination Property in Multicriteria Optimization*, Journal of Mathematical Analysis and Applications **114**, No. 1 (1986), 7–16.

Henig M.I., *Value Functions, Domination Cones and Proper Efficiency in Multicriteria Optimization*, Mathematical Programming **46**, No. 2 (1990), 205–217.

Henig M.I., Buchanan J.T., *Generalized Tradeoff Directions in Multiobjective Optimization Problems*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 47–56.

Henig M.I., Buchanan J.T., *Tradeoff Directions in Multiobjective Optimization Problems*, Mathematical Programming **78**, No. 3 (1997), 357–374.

Hobbs B.F., *What Can We Learn From Experiments in Multiobjective Decision Analysis?*, IEEE Transactions on Systems, Man, and Cybernetics **16**, No. 3 (1986), 384–394.

Hobbs B.F., Chankong V., Hamadeh W., Stakhiv E.Z., *Does Choice of Multicriteria Method Matter? An Experiment in Water Resources Planning*, Water Resources Research **28**, No. 7 (1992), 1767–1779.

Horst R., Thoai N.V., *Utility Function Programs and Optimization over the Efficient Set in Multiple-Objective Decision Making*, Journal of Optimization Theory and Applications **92**, No. 3 (1997), 605–631.

Hu Y., *Theory and Application of Multiobjective Optimization in China – A Survey*, Annals of Operations Research **24**, No. 1–4 (1990), 45–51.

Hussein M.L., El-Ghaffar F.S.A., *An Interactive Approach for Vector Optimization Problems*, European Journal of Operational Research **89**, No. 1 (1996), 185–192.

Hwang C.-L., Lin M.-J., *Group Decision Making under Multiple Criteria: Methods and Applications*, Lecture Notes in Economics and Mathematical Systems 281, Springer-Verlag, New York, 1987.

Hwang C.-L., Masud A.S.M., *Multiple Objective Decision Making – Methods and Applications: A State-of-the-Art Survey*, Lecture Notes in Economics and Mathematical Systems 164, Springer-Verlag, Berlin, Heidelberg, 1979.

Hwang C.L., Yoon K., *Multiple Attribute Decision Making Methods and Applications: A State-of-the-Art Survey*, Lecture Notes in Economics and Mathematical Systems 186, Springer-Verlag, Berlin, Heidelberg, 1981.

Hwang C.L., Paidy S.R., Yoon K., Masud A.S.M., *Mathematical Programming with Multiple Objectives: A Tutorial*, Computers & Operations Research **7** (1980), 5–31.

Ignizio J.P., *Goal Programming and Extensions*, Lexington Books, D.C. Heath and Company, 1976.

Ignizio J.P., *Generalized Goal Programming: An Overview*, Computers & Operations Research **10**, No. 4 (1983a), 277–289.

Ignizio J.P., *An Approach to the Modeling and Analysis of Multiobjective Generalized Networks*, European Journal of Operational Research **12**, No. 4 (1983b), 357–361.

Ignizio J.P., *Introduction to Linear Goal Programming*, Sage Publications, Sage University Press, Beverly Hills, 1985.

Inuiguchi M., Kume Y., *Goal Programming Problems with Interval Coefficients and Target Intervals*, European Journal of Operational Research **52**, No. 3 (1991), 345–360.

Isermann H., *Proper Efficiency and the Linear Vector Maximization Problem*, Operations Research **22**, No. 1 (1974), 189–191.

Isermann H., Steuer R.E., *Computational Experience Concerning Payoff Tables and Minimum Criterion Values over the Efficient Set*, European Journal of Operational Research **33**, No. 1 (1988), 91–97.

Ishizuka Y., Shimizu K., *Necessary and Sufficient Conditions for the Efficient Solutions of Nondifferentiable Multiobjective Problems*, IEEE Transactions on Systems, Man, and Cybernetics **14**, No. 4 (1984), 625–629.

Jahn J., *Duality in Vector Optimization*, Mathematical Programming **25**, No. 3 (1983), 343–353.

Jahn J., *Scalarization in Vector Optimization*, Mathematical Programming **29**, No. 2 (1984), 203–218.

Jahn J., *Scalarization in Multi Objective Optimization*, Mathematics of Multi Objective Optimization, Edited by P. Serafini, Springer-Verlag, Wien, New York, 1985, pp. 45–88.

Jahn J., *Mathematical Vector Optimization in Partially Ordered Linear Spaces*, Methoden und Verfahren der mathematischen Physik, Band 31, Verlag Peter Lang GmbH, Frankfurt am Main, 1986a.

Jahn J., *Existence Theorems in Vector Optimization*, Journal of Optimization Theory and Applications **50**, No. 3 (1986b), 397–406.

Jahn J., *Parametric Approximation Problems Arising in Vector Optimization*, Journal of Optimization Theory and Applications **54**, No. 3 (1987), 503–516.

Janssen R., *Multiobjective Decision Support for Environmental Management*, Kluwer Academic Publishers, Dordrecht, 1992.

Jarvenpaa S.L., *The Effect of Task Demands and Graphical Format on Information Processing Strategies*, Management Science **35**, No. 3 (1989), 285–303.

Jaszkiewicz A., Slowiński R., *The Light Beam Search over a Non-Dominated Surface of a Multiple-Objective Programming Problem*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 87–99.

Jaszkiewicz A., Slowiński R., *The Light Beam Search – Outranking Based Interactive Procedure for Multiple-Objective Mathematical Programming*, Advances in Multicriteria Analysis, Edited by P.M. Pardalos, Y. Siskos, C. Zopounidis, Kluwer Academic Publishers, Dordrecht, 1995, pp. 129–146.

Jedrzejowicz P., Rosicka L., *Multicriterial Reliability Optimization Problem*, Foundations of Control Engineering **8**, No. 3–4 (1983), 165–173.

Jelassi M.T., Ozernoy V.M., *A Framework for Building an Expert System for MCDM Models Selection*, Improving Decision Making in Organisations, Edited by A.G. Lockett, G. Islei, Lecture Notes in Economics and Mathematical Systems 335, Springer-Verlag, Berlin, Heidelberg, 1989, pp. 553–562.

Jelassi M.T., Jarke M., Stohr E.A., *Designing a Generalized Multiple Criteria Decision Support System*, Decision Making with Multiple Objectives, Edited by Y.Y. Haimes, V. Chankong, Lecture Notes in Economics and Mathematical Systems 242, Springer-Verlag, 1985, pp. 214–235.

Jendo S., *Multicriteria Optimization of Single-Layer Cable System*, Archives of Mechanics **38**, No. 3 (1986), 219–234.

Jeyakumar V., Yang X.Q., *Convex Composite Multi-Objective Nonsmooth Programming*, Mathematical Programming **59**, No. 3 (1993), 325–343.

Jurkiewicz E., *Stability of Compromise Solution in Multicriteria Decision-Making Problems*, Journal of Optimization Theory and Applications **40**, No. 1 (1983), 77–83.

Kacprzyk J., Orlovski S.A. (Eds.), *Optimization Models Using Fuzzy Sets and Possibility Theory*, Theory and Decision Library, Series B: Mathematical and Statistical Methods, D. Reidel Publishing Company, 1987.

Kaliszewski I., *A Characterization of Properly Efficient Solutions by an Augmented Tchebycheff Norm*, Bulletin of the Polish Academy of Sciences – Technical Sciences **33**, No. 7–8 (1985), 415–420.

Kaliszewski I., *Norm Scalarization and Proper Efficiency in Vector Optimization*, Foundations of Control Engineering **11**, No. 3 (1986), 117–131.

Kaliszewski I., *A Modified Weighted Tchebycheff Metric for Multiple Objective Programming*, Computers & Operations Research **14**, No. 4 (1987), 315–323.

Kaliszewski I., *Quantitative Pareto Analysis by Cone Separation Technique*, Kluwer Academic Publishers, 1994.

Kaliszewski I., *A Theorem on Nonconvex Functions and Its Application to Vector Optimization*, European Journal of Operational Research **80**, No. 2 (1995), 439–445.

Kaliszewski I., Michalowski W., *Generation of Outcomes with Selectively Bounded Trade-Offs*, Foundations of Computing and Decision Sciences **20**, No. 2 (1995), 113–122.

Kaliszewski I., Michalowski W., *Efficient Solutions and Bounds on Tradeoffs*, Journal of Optimization Theory and Applications **94**, No. 2 (1997), 381–394.

Kaliszewski I., Michalowski W., Kersten G., *A Hybrid Interactive Technique for the MCDM Problems*, Essays in Decision Making: A Volume in Honour of Stanley Zionts, Edited by M.H. Karwan, J. Spronk, J. Wallenius, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 48–59.

Kasanen E., Östermark R., Zeleny M., *Gestalt System of Holistic Graphics: New Management Support View of MCDM*, Computers & Operations Research **18**, No. 2 (1991), 233–239.

Kaul R.N., Lyall V., Kaur S., *Semilocal Pseudolinearity and Efficiency*, European Journal of Operational Research **36**, No. 3 (1988), 402–409.

Keeney R.L., Raiffa H., *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley & Sons, Inc., 1976.

Kim S.H., Gal T., *A New Interactive Algorithm for Multi-Objective Linear Programming Using Maximally Changeable Dominance Cone*, European Journal of Operational Research **64**, No. 1 (1993), 126–137.

Kim S.-H., Ahn B.-S., Choi S.-H., *An Efficient Force Planning System Using Multi-Objective Linear Goal Programming*, Computers & Operations Research **24**, No. 6 (1997), 569–580.

Kirilov L.M., Vassilev V.S., *A Method for Solving Multiple Objective Linear Programming Problems*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 302–309.

Kitagawa H., Watanabe N., Nishimura Y., Matsubara M., *Some Pathological Configurations of Noninferior Set Appearing in Multicriteria Optimization Problems of Chemical Processes*, Journal of Optimization Theory and Applications **38**, No. 4 (1982), 541–563.

Kiwiel K.C., *An Aggregate Subgradient Descent Method for Solving Large Convex Nonsmooth Multiobjective Minimization Problems*, Large Scale Systems: Theory and Applications 1983, Edited by A. Straszak, Pergamon Press, 1984, pp. 283–288.

Kiwiel K.C., *A Descent Method for Nonsmooth Convex Multiobjective Minimization*, Large Scale Systems **8**, No. 2 (1985a), 119–129.

Kiwiel K.C., *An Algorithm for Linearly Constrained Nonsmooth Convex Multiobjective Minimization*, Systems Analysis and Simulation 1985 Part I: Theory and Foundations, Edited by A. Sydow, S.M. Thoma, R. Vichnevetsky, Akademie-Verlag, Pergamon Press, Berlin, 1985b, pp. 236–238.

Kiwiel K.C., *Methods of Descent for Nondifferentiable Optimization*, Lecture Notes in Mathematics 1133, Springer-Verlag, Berlin, Heidelberg, 1985c.

Kiwiel K.C., *A Method for Solving Certain Quadratic Programming Problems Arising in Nonsmooth Optimization*, IMA Journal of Numerical Analysis **6**, No. 2 (1986), 137–152.

Kiwiel K.C., *Proximity Control in Bundle Methods for Convex Nondifferentiable Minimization*, Mathematical Programming **46**, No. 1 (1990), 105–122.

Kiwiel K.C., Stachurski A., *Issues of Effectiveness Arising in the Design of a System of Nondifferentiable Optimization Algorithms*, Aspiration Based Decision Support Systems: Theory, Software and Applications, Edited by A. Lewandowski, A. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 331, Springer-Verlag, 1989, pp. 180–192.

Klimberg R., *GRADS: A New Graphical Display System for Visualizing Multiple Criteria Solutions*, Computers & Operations Research **19**, No. 7 (1992), 707–711.

Klinger A., *Improper Solutions of the Vector Maximum Problem*, Operations Research **15**, No. 3 (1967), 570–572.

Kok M., *Scalarization and the Interface with Decision Makers in Interactive Multi Objective Linear Programming*, Mathematics of Multi Objective Optimization, Edited by P. Serafini, Springer-Verlag, Wien, New York, 1985, pp. 433–438.

Kok M., *The Interface with Decision Makers and Some Experimental Results in Interactive Multiple Objective Programming Method*, European Journal of Operational Research **26**, No. 1 (1986), 96–107.

Kok M., Lootsma F.A., *Pairwise-Comparison Methods in Multiple Objective Programming, with Applications in a Long-Term Energy-Planning Model*, European Journal of Operational Research **22**, No. 1 (1985), 44–55.

Köksalan M.M., Moskowitz H., *Solving the Multiobjective Decision Making Problem Using a Distance Function*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 101–107.

Koopmans T.C., *Analysis and Production as an Efficient Combination of Activities*, Activity Analysis of Production and Allocation, Edited by T.C. Koopmans, Yale University Press, New Haven, London, 1971 (originally published in 1951), pp. 33–97.

Korhonen P.J., *Solving Discrete Multiple Criteria Problems by Using Visual Interaction*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986, pp. 176–185.

Korhonen P.J., *VIG (A Visual Interactive Approach to Goal Programming), User's Guide*, NumPlan (1987).

Korhonen P.J., *The Multiobjective Linear Programming Decision Support System VIG and Its Applications*, Readings in Multiple Criteria Decision Aid, Edited by C.A. Bana e Costa, Springer-Verlag, Berlin, Heidelberg, 1990, pp. 471–491.

Korhonen P.J., *Two Decision Support Systems for Continuous and Discrete Multiple Criteria Decision Making: VIG and VIMDA*, Methodology, Implementation and Applications of Decision Support Systems, Edited by A. Lewandowski, P. Serafini, M.G. Speranza, International Centre for Mechanical Sciences, Courses and Lectures No. 320, Springer-Verlag, Wien, New York, 1991a, pp. 85–103.

Korhonen P.J., *Using Harmonious Houses for Visual Pairwise Comparison of Multiple Criteria Alternatives*, Decision Support Systems **7**, No. 1 (1991b), 47–54.

Korhonen P.J., *Reference Direction Approach to Multiple Objective Linear Programming: Historical Overview*, Essays in Decision Making: A Volume in Honour of Stanley Zionts, Edited by M.H. Karwan, J. Spronk, J. Wallenius, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 74–92.

Korhonen P., Halme M., *Using Lexicographic Parametric Programming for Searching a Nondominated Set in Multiple Objective Linear Programming*, Journal of Multi-Criteria Decision Analysis **5**, No. 4 (1996), 291–300.

Korhonen P., Laakso J., *A Visual Interactive Method for Solving the Multiple-Criteria Problem*, Interactive Decision Analysis, Edited by M. Grauer, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 229, Springer-Verlag, 1984, pp. 146–153.

Korhonen P., Laakso J., *On Developing a Visual Interactive Multiple Criteria Method – An Outline*, Decision Making with Multiple Objectives, Edited by Y.Y. Haimes, V. Chankong, Lecture Notes in Economics and Mathematical Systems 242, Springer-Verlag, Berlin, Heidelberg, 1985, pp. 272–281.

Korhonen P., Laakso J., *A Visual Interactive Method for Solving the Multiple Criteria Problem*, European Journal of Operational Research **24**, No. 2 (1986a), 277–287.

Korhonen P., Laakso J., *Solving Generalized Goal Programming Problems Using a Visual Interactive Approach*, European Journal of Operational Research **26**, No. 3 (1986b), 355–363.

Korhonen P., Narula S.C., *An Evolutionary Approach to Support Decision Making with Linear Decision Models*, Journal of Multi-Criteria Decision Analysis **2**, No. 2 (1993), 111–119.

Korhonen P., Wallenius J., *A Pareto Race*, Naval Research Logistics **35** (1988), 615–623.

Korhonen P., Wallenius J., *A Careful Look at Efficiency and Utility in Multiple Criteria Decision Making: A Tutorial*, Asia–Pacific Journal of Operational Research **6**, No. 1 (1989a), 46–62.

Korhonen P., Wallenius J., *Observations Regarding Choice Behaviour in Interactive Multiple Criteria Decision-Making Environments: An Experimental Investigation*, Methodology and Software for Interactive Decision Support, Edited by A. Lewandowski, I. Stanchev, Lecture Notes in Economics and Mathematical Systems 337, Springer-Verlag, 1989b, pp. 163–170.

Korhonen P., Wallenius J., *VIG – A Visual and Dynamic Decision Support System for Multiple Objective Linear Programming*, Multiple Criteria Decision Making and Risk Analysis Using Microcomputers, Edited by B. Karpak, S. Zionts, Springer-Verlag, Berlin, Heidelberg, 1989c, pp. 251–281.

Korhonen P., Wallenius J., *A Multiple Objective Linear Programming Decision Support System*, Decision Support Systems **6**, No. 3 (1990), 243–251.

Korhonen P., Wallenius J., *Behavioural Issues in MCDM: Neglected Research Questions*, Journal of Multi-Criteria Decision Analysis **5**, No. 3 (1996), 178–182.

Korhonen P., Wallenius J., *Behavioral Issues in MCDM: Neglected Research Questions*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 412–422.

Korhonen P., Wallenius J., Zionts S., *Solving the Discrete Multiple Criteria Problem Using Convex Cones*, Management Science **30**, No. 11 (1984), 1336–1345.

Korhonen P., Moskowitz H., Wallenius J., *Choice Behavior in Interactive Multiple-Criteria Decision Making*, Annals of Operations Research **23**, No. 1–4 (1990), 161–179.

Korhonen P., Moskowitz H., Wallenius J., *Multiple Criteria Decision Support – A Review*, European Journal of Operational Research **63**, No. 3 (1992a), 361–375.

Korhonen P., Wallenius J., Zionts S., *A Computer Graphics-Based Decision Support System for Multiple Objective Linear Programming*, European Journal of Operational Research **60**, No. 3 (1992b), 280–286.

Korhonen P., Salo S., Steuer R.E., *A Heuristic for Estimating Nadir Criterion Values in Multiple Objective Linear Programming*, Operations Research **45**, No. 5 (1997), 751–757.

Kornbluth J.S.H., *A Survey of Goal Programming*, OMEGA **1**, No. 2 (1973), 193–205.

Koski J., Silvennoinen R., *Norm Methods and Partial Weighting in Multicriterion Optimization of Structures*, International Journal for Numerical Methods in Engineering **24**, No. 6 (1987), 1101–1121.

Kostreva M.M., Ordoyne T.J., Wiecek M., *Multiple-Objective Programming with Polynomial Objectives and Constraints*, European Journal of Operational Research **57**, No. 3 (1992), 381–394.

Kostreva M.M., Zheng Q., Zhuang D., *A Method for Approximating Solutions of Multicriterial Nonlinear Optimization Problems*, Optimization Methods and Software **5**, No. 3 (1995), 209–226.

Kreglewski T., *Nonlinear Optimization Techniques in Decision Support Systems*, Aspiration Based Decision Support Systems: Theory, Software and Applications, Edited by A. Lewandowski, A. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 331, Springer-Verlag, 1989, pp. 158–171.

Kreglewski T., Paczynski J., Wierzbicki A.P., *IAC-DIDAS-N A Dynamic Interactive Decision Analysis and Support System for Multicriteria Analysis of Nonlinear Models on Professional Microcomputers*, Theory, Software and Testing Examples for Decision Support Systems, Edited by A. Lewandowski, A. Wierzbicki, Working Paper WP-87-26, IIASA, Laxenburg, Austria, 1987, pp. 177–192.

Kreglewski T., Granat J., Wierzbicki A.P., *IAS-DIDAS-N: A Dynamic Interactive Decision Analysis and Support System for Multicriteria Analysis of Nonlinear Models, v. 4.0*, Collaborative Paper CP-91-010, IIASA, Laxenburg (1991).

Kuhn H.W., Tucker A.W., *Nonlinear Programming*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Edited by J. Neyman, University of California Press, Berkeley, Los Angeles, 1951, pp. 481–492.

Kuk H., Tanino T., Tanaka M., *Sensitivity Analysis in Parametrized Convex Vector Optimization*, Journal of Mathematical Analysis and Applications **202**, No. 2 (1996), 511–522.

Kumar P., Singh N., Tewari N.K., *A Nonlinear Goal Programming Model for Multistage, Multiobjective Decision Problems with Application to Grouping and Loading Problem in a Flexible Manufacturing System*, European Journal of Operational Research **53**, No. 2 (1991), 166–171.

Larichev O.I., Polyakov O.A., Nikiforov A.O., *Multicriterion Linear Programming Problems*, Journal of Economic Psychology **8** (1987), 389–407.

Lazimy R., *Solving Multiple Criteria Problems by Interactive Decomposition*, Mathematical Programming **35**, No. 3 (1986a), 334–361.

Lazimy R., *Interactive Relaxation Method for a Broad Class of Integer and Continuous Nonlinear Multiple Criteria Problems*, Journal of Mathematical Analysis and Applications **116**, No. 2 (1986b), 553–573.

Lee G.M., *On Efficiency in Nonlinear Fractional Vector Maximization Problem*, Optimization **25**, No. 1 (1992), 47–52.

Levary R.R., *Optimal Control Problems with Multiple Goal Objectives*, Optimal Control Applications & Methods **7**, No. 2 (1986), 201–207.

Lewandowski A., *Man-Machine Dialogue Interfaces in Decision Support Systems*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986, pp. 161–175.

Lewandowski A., Granat J., *Dynamic BIPLOT as the Interaction Interface for Aspiration Based Decision Support Systems*, Multiple Criteria Decision Support, Edited by P. Korhonen, A. Lewandowski, J. Wallenius, Lecture Notes in Economics and Mathematical Systems 356, Springer-Verlag, 1991, pp. 229–241.

Lewandowski A., Grauer M., *The Reference Point Optimization Approach – Methods of Efficient Implementation*, Working Paper WP-82-26, IIASA, Laxenburg (1982).

Lewandowski A., Rogowski T., Kreglewski T., *A Trajectory-Oriented Extension of DIDASS and Its Applications*, Plural Rationality and Interactive Decision Processes, Edited by M. Grauer, M. Thompson, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 248, Springer-Verlag, Berlin, Heidelberg, 1985a, pp. 261–268.

Lewandowski A., Rogowski T., Kreglewski T., *Application of DIDASS Methodology to a Flood Control Problem – Numerical Experiments*, Plural Rationality and Interactive Decision Processes, Edited by M. Grauer, M. Thompson, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 248, Springer-Verlag, Berlin, Heidelberg, 1985b, pp. 325–328.

Lewandowski A., Kreglewski T., Rogowski T., Wierzbicki A.P., *Decision Support Systems of DIDAS Family (Dynamic Interactive Decision Analysis & Support)*, Theory, Software and Testing Examples for Decision Support Systems, Edited by A. Lewandowski, A. Wierzbicki, Working Paper WP-87-26, IIASA, Laxenburg, 1987, pp. 4–26.

Li D., *Convexification of a Noninferior Frontier*, Journal of Optimization Theory and Applications **88**, No. 1 (1996), 177–196.

Li D., Haimes Y.Y., *The Envelope Approach for Multiobjective Optimization Problems*, IEEE Transactions on Systems, Man, and Cybernetics **17**, No. 6 (1987), 1026–1038; *Correction*, IEEE Transactions on Systems, Man, and Cybernetics **18**, No. 2 (1988), 332.

Lieberman E.R., *Soviet Multi-Objective Programming Methods: An Overview*, Multiobjective Problems of Mathematical Programming, Edited by A. Lewandowski, V. Volkovich, Lecture Notes in Economics and Mathematical Systems 351, Springer-Verlag, 1991a, pp. 21–31.

Lieberman E.R., *Soviet Multi-Objective Mathematical Programming Methods: An Overview*, Management Science **37**, No. 9 (1991b), 1147–1165.

Lin J.G., *Multiple-Objective Problems: Pareto-Optimal Solutions by Methods of Proper Equality Constraints*, IEEE Transactions on Automatic Control **21**, No. 5 (1976a), 641–650.

Lin J.G., *Three Methods for Determining Pareto-Optimal Solutions of Multiple-Objective Problems*, Directions in Large-Scale Systems, Many-Person Optimization and Decentralized Control, Edited by Y.C. Ho, S.K. Mitter, Plenum Press, New York, 1976b, pp. 117–138.

Liu J.C., *ϵ-Pareto Optimality for Nondifferentiable Multiobjective Programming via Penalty Function*, Journal of Mathematical Analysis and Applications **198**, No. 1 (1996), 248–261.

Liu L., *The Second-Order Conditions of Nondominated Solutions for $C^{1,1}$ Generalized Multiobjective Mathematical Programming*, Systems Science and Mathematical Sciences **4**, No. 2 (1991), 128–138.

Liu L., Křížek M., *The Second Order Optimality Conditions for Nonlinear Mathematical Programming with $C^{1,1}$ Data*, Applications of Mathematics **42**, No. 4 (1997), 311–320.

Loganathan G.V., Sherali H.D., *A Convergent Interactive Cutting-Plane Algorithm for Multiobjective Optimization*, Operation Research **35**, No. 3 (1987), 365–377.

Lotfi V., Teich J.E., *Multiple Criteria Decision Making Using Personal Computers*, Working Paper No. 727, State University of New York at Buffalo, Buffalo, New York (1989).

Lotov A.V., *An Estimate of Solution Set Perturbations for a System of Linear Inequalities*, Optimization Methods and Software **6**, No. 1 (1995), 1–24.

Lotov A.V., *Comment on D.J. White "A Characterization of the Feasible Set of Objective Function Vectors in Linear Multiple Objective Problems"*, European Journal of Operational Research **89**, No. 1 (1996), 215–220.

Lotov A., Chernykh O., Hellman O., *Multiple Objective Analysis of Long-Term Development Strategies for a National Economy*, European Journal of Operational Research **56**, No. 2 (1992), 210–218.

Lotov A.V., Bushenkov V.A., Chernykh O.L., *Multicriteria DSS for River Water-Quality Planning*, Microcomputers in Civil Engineering **12** (1997), 57–67.

Lowe T.J., Thisse J.-F., Ward J.E., Wendell R.E., *On Efficient Solutions to Multiple Objective Mathematical Programs*, Management Science **30**, No. 11 (1984), 1346–1349.

Luc D.T., *On the Domination Property in Vector Optimization*, Journal of Optimization Theory and Applications **43**, No. 2 (1984a), 327–330.

Luc D.T., *On Duality Theory in Multiobjective Programming*, Journal of Optimization Theory and Applications **43**, No. 4 (1984b), 557–582.

Luc D.T., *On Scalarizing Method in Vector Optimization*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986, pp. 46–51.

Luc D.T., *About Duality and Alternative in Multiobjective Optimization*, Journal of Optimization Theory and Applications **53**, No. 2 (1987), 303–307.

Luc D.T., *Theory of Vector Optimization*, Lecture Notes in Economics and Mathematical Systems 319, Springer-Verlag, Berlin, Heidelberg, 1989.

Luc D.T., *Recession Cones and the Domination Property in Vector Optimization*, Mathematical Programming **49**, No. 1 (1990), 113–122.

Luc D.T., *On the Properly Efficient Points of Nonconvex Sets*, European Journal of Operational Research **86**, No. 2 (1995), 262–269.

Luc D.T., Jahn J., *Axiomatic Approach to Duality in Optimization*, Numerical Functional Analysis and Optimization **13**, No. 3-4 (1992), 305–326.

Luc D.T., Schaible S., *Efficiency and Generalized Concavity*, Journal of Optimization Theory and Applications **94**, No. 1 (1997), 147–153.

Lucchetti R., *Some Existence Results and Stability in Multi Objective Optimization*, Mathematics of Multi Objective Optimization, Edited by P. Serafini, Springer-Verlag, Wien, New York, 1985, pp. 179–187.

Luenberger D.G., *Linear and Nonlinear Programming*, Second Edition, Addison-Wesley Publishing Company, Inc., 1984.

MacCrimmon K.R., *An Overview of Multiple Objective Decision Making*, Multiple Criteria Decision Making, Edited by J.L. Cochrane, M. Zeleny, University of South Carolina Press, Columbia, South Carolina, 1973, pp. 18–44.

Majumdar A.A.K., *Optimality Conditions in Differentiable Multiobjective Programming*, Journal of Optimization Theory and Applications **92**, No. 2 (1997), 419–427.

Mäkelä M.M., *Issues of Implementing a Fortran Subroutine Package NSOLIB for Nonsmooth Optimization*, Report 5/1993, University of Jyväskylä, Department of Mathematics, Laboratory of Scientific Computing, Jyväskylä (1993).

Mäkelä M.M., Neittaanmäki P., *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*, World Scientific Publishing Co., Singapore, 1992.

Mañas M., *Graphical Methods of Multicriterial Optimization*, Zeitschrift für Angewandte Mathematik und Mechanik **62**, No. 5 (1982), 375–377.

Mangasarian O.L., *Nonlinear Programming*, McGraw-Hill, Inc., 1969.

Mareschal B., Brans J.-P., *Geometrical Representation for MCDA*, European Journal of Operational Research **34**, No. 1 (1988), 69–77.

Martein L., *Some Results on Regularity in Vector Optimization*, Optimization **20**, No. 6 (1989), 787–798.

Martel J.-M., Aouni B., *Incorporating the Decision-Maker's Preferences in the Goal-Programming Model*, Journal of the Operational Research Society **41**, No. 12 (1990), 1121–1132.

Martel J.-M., Aouni B., *Diverse Imprecise Goal Programming Model Formulations*, Journal of Global Optimization **12**, No. 2 (1998), 127–138.

Martínez-Legaz J.-E., *Lexicographical Order and Duality in Multiobjective Programming*, European Journal of Operational Research **33**, No. 3 (1988), 342–348.

Martínez-Legaz J.-E., Singer I., *Surrogate Duality for Vector Optimization*, Numerical Functional Analysis and Optimization **9**, No. 5–6 (1987), 547–568.

Marusciac I., *On Fritz John Type Optimality Criterion in Multi-Objective Optimization*, Revue d'Analyse Numérique et de Théorie de l'Approximation **11**, No. 1–2 (1982), 109–114.

Masud A.S.M., Hwang C.L., *Interactive Sequential Goal Programming*, Journal of the Operational Research Society **32** (1981), 391–400.

Masud A.S.M., Zheng X., *An Algorithm for Multiple-Objective Non-Linear Programming*, Journal of the Operational Research Society **40**, No. 10 (1989), 895–906.

Matos M.A., Borges P., *A Flexible Interface for Decision-Aid in Multicriteria Decision Problems*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 390–400.

Meisel W.S., *Tradeoff Decision in Multiple Criteria Decision Making*, Multiple Criteria Decision Making, Edited by J.L. Cochrane, M. Zeleny, University of South Carolina Press, Columbia, South Carolina, 1973, pp. 461–476.

Michalowski W., *Evaluation of a Multiple Criteria Interactive Programming Approach: An Experiment*, INFOR **25**, No. 2 (1987), 165–173.

Michalowski W., *MCDM at the Crossroads*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 579–584.

Michalowski W., Szapiro T., *A Bi-Reference Procedure for Interactive Multiple Criteria Programming*, Operations Research **40**, No. 2 (1992), 247–258.

Miettinen K., *On the Methodology of Multiobjective Optimization with Applications*, Doctoral Thesis, Report 60, University of Jyväskylä, Department of Mathematics, Jyväskylä, 1994.

Miettinen K., Mäkelä M.M., *Nonsmooth Multicriteria Optimization Applied to Optimal Control*, Reports on Applied Mathematics and Computing, No. 6, University of Jyväskylä, Jyväskylä (1991).

Miettinen K., Mäkelä M.M., *An Interactive Method for Nonsmooth Multiobjective Optimization with an Application to Optimal Control*, Optimization Methods and Software **2** (1993), 31–44.

Miettinen K., Mäkelä M.M., *A Nondifferentiable Multiple Criteria Optimization Method Applied to Continuous Casting Process*, Proceedings of the Seventh European Conference on Mathematics in Industry, Edited by A. Fasano, M. Primicerio, B.G. Teubner, Stuttgart, 1994, pp. 255–262.

Miettinen K., Mäkelä M.M., *Interactive Bundle-Based Method for Nondifferentiable Multiobjective Optimization: NIMBUS*, Optimization **34**, No. 3 (1995), 231–246.

Miettinen K., Mäkelä M.M., *NIMBUS – Interactive Method for Nondifferentiable Multiobjective Optimization Problems*, Multi-Objective Programming and Goal Programming: Theories and Applications, Edited by M. Tamiz, Lecture Notes in Economics and Mathematical Systems 432, Springer-Verlag, Berlin, Heidelberg, 1996a, pp. 50–57.

Miettinen K., Mäkelä M.M., *Comparing Two Versions of NIMBUS Optimization System*, Report 23/1996, University of Jyväskylä, Department of Mathematics, Laboratory of Scientific Computing, Jyväskylä, 1996b.

Miettinen K., Mäkelä M.M., *Interactive Method NIMBUS for Nondifferentiable Multiobjective Optimization Problems*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 310–319.

Miettinen K., Mäkelä M.M., *Theoretical and Computational Comparison of Multiobjective Optimization Methods NIMBUS and RD*, Report 5/1998, University of Jyväskylä, Department of Mathematics, Laboratory of Scientific Computing, Jyväskylä, 1998a.

Miettinen K., Mäkelä M.M., *Interactive MCDM Support System in the Internet*, Trends in Multicriteria Decision Making: Proceedings of the 13th International Conference on Multiple Criteria Decision Making, Edited by T. Stewart, R. van den Honert, Springer-Verlag, 1998b, pp. 419–428.

Miettinen K., Mäkelä M.M., Mäkinen R.A.E., *Interactive Multiobjective Optimization System NIMBUS Applied to Nonsmooth Structural Design Problems*, System Modelling and Optimization, Proceedings of the 17th IFIP Conference on System Modelling and Optimization, Prague, Czech Republic, Edited by J. Doležal, J. Fidler, Chapman & Hall, London, 1996a, pp. 379–385.

Miettinen K., Mäkelä M.M., Männikkö T., *Nondifferentiable Multiobjective Optimizer NIMBUS Applied to an Optimal Control Problem of Continuous Casting*, Report 22/1996, University of Jyväskylä, Department of Mathematics, Laboratory of Scientific Computing, Jyväskylä, 1996b.

Miller G.A., *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*, Psychological Review **63**, No. 2 (1956), 81–87.

Minami M., *Weak Pareto Optimality of Multiobjective Problem in a Banach Space*, Bulletin of Mathematical Statistics **19**, No. 3–4 (1980–81), 19–23.

Minami M., *Weak Pareto Optimality of Multiobjective Problems in a Locally Convex Linear Topological Space*, Journal of Optimization Theory and Applications **34**, No. 4 (1981), 469–484.

Minami M., *Weak Pareto-Optimal Necessary Conditions in a Nondifferentiable Multiobjective Program on a Banach Space*, Journal of Optimization Theory and Applications **41**, No. 3 (1983), 451–461.

Mishra S.K., *Lagrange Multipliers Saddle Point and Scalarizations in Composite Multiobjective Nonsmooth Programming*, Optimization **38**, No. 2 (1996), 93–105.

Mishra S.K., Mukherjee R.N., *Generalized Convex Composite Multi-Objective Nonsmooth Programming and Conditional Proper Efficiency*, Optimization **34**, No. 1 (1995), 53–66.

Mitani K., Nakayama H., *A Multiobjective Diet Planning Support System Using the Satisficing Trade-Off Method*, Journal of Multi-Criteria Decision Analysis **6**, No. 3 (1997), 131–139.

Mitra A., Patankar J.G., *A Multi-Objective Model for Warranty Estimation*, European Journal of Operational Research **45**, No. 2–3 (1990), 347–355.

Miyaji I., Ohno K., Mine H., *Solution Method for Partitioning Students into Groups*, European Journal of Operational Research **33**, No. 1 (1988), 82–90.

Mocci U., Primicerio L., *Ring Network Design: An MCDM Approach*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 491–500.

Moldavskiy M.A., *Singling Out a Set of Undominated Solutions in Continuous Vector Optimization Problems*, Soviet Automatic Control **14**, No. 5 (1981), 47–53.

Monarchi D.E., Kisiel C.C., Duckstein L., *Interactive Multiobjective Programming in Water Resources: A Case Study*, Water Resources Research **9**, No. 4 (1973), 837–850.

Monarchi D.E., Weber J.E., Duckstein L., *An Interactive Multiple Objective Decision-Making Aid Using Nonlinear Goal Programming*, Multiple Criteria Decision Making Kyoto 1975, Edited by M. Zeleny, Lecture Notes in Economics and Mathematical Systems 123, Springer-Verlag, Berlin, Heidelberg, 1976, pp. 235–253.

Morón M.A., Romero C., Ruiz del Portal F.R., *Generating Well-Behaved Utility Functions for Compromise Programming*, Journal of Optimization Theory and Applications **91**, No. 3 (1996), 643–649.

Mote J., Olson D.L., Venkataramanan M.A., *A Comparative Multiobjective Programming Study*, Mathematical and Computer Modelling **10**, No. 10 (1988), 719–729.

M'silti A., Tolla P., *An Interactive Multiobjective Nonlinear Programming Procedure*, European Journal of Operational Research **64**, No. 1 (1993), 115–125.

Mukai H., *Algorithms for Multicriterion Optimization*, IEEE Transactions on Automatic Control **25**, No. 2 (1980), 177–186.

Musselman K., Talavage J., *A Tradeoff Cut Approach to Multiple Objective Optimization*, Operations Research **28**, No. 6 (1980), 1424–1435.

Nakayama H., *Geometric Consideration of Duality in Vector Optimization*, Journal of Optimization Theory and Applications **44**, No. 4 (1984), 625–655.

Nakayama H., *On the Components in Interactive Multiobjective Programming Methods*, Plural Rationality and Interactive Decision Processes, Edited by M. Grauer, M. Thompson, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 248, Springer-Verlag, Berlin, Heidelberg, 1985a, pp. 234–247.

Nakayama H., *Lagrange Duality and Its Geometric Interpretation*, Mathematics of Multi Objective Optimization, Edited by P. Serafini, Springer-Verlag, Wien, New York, 1985b, pp. 105–127.

Nakayama H., *Duality Theory in Vector Optimization: An Overview*, Decision Making with Multiple Objectives, Edited by Y.Y. Haimes, V. Chankong, Lecture Notes in Economics and Mathematical Systems 242, Springer-Verlag, 1985c, pp. 109–125.

Nakayama H., *Sensitivity and Trade-Off Analysis in Multiobjective Programming*, Methodology and Software for Interactive Decision Support, Edited by A. Lewandowski, I. Stanchev, Lecture Notes in Economics and Mathematical Systems 337, Springer-Verlag, 1989, pp. 86–93.

Nakayama H., *Satisficing Trade-Off Method for Problems with Multiple Linear Fractional Objectives and Its Applications*, Multiobjective Problems of Mathematical Programming, Edited by A. Lewandowski, V. Volkovich, Lecture Notes in Economics and Mathematical Systems 351, Springer-Verlag, 1991a, pp. 42–50.

Nakayama H., *Trade-off Analysis Based upon Parametric Optimization*, Multiple Criteria Decision Support, Edited by P. Korhonen, A. Lewandowski, J. Wallenius, Lecture Notes in Economics and Mathematical Systems 356, Springer-Verlag, 1991b, pp. 42–52.

Nakayama H., *Trade-Off Analysis Using Parametric Optimization Techniques*, European Journal of Operational Research **60**, No. 1 (1992a), 87–98.

Nakayama H., *Theoretical Remarks on Dynamic Trade-Off*, Multiple Criteria Decision Making: Proceedings of the Ninth International Conference: Theory and Applications in Business, Industry, and Government, Edited by A. Goicoechea, L. Duckstein, S. Zionts, Springer-Verlag, New York, 1992b, pp. 297–309.

Nakayama H., *Engineering Applications of Multi-Objective Programming: Recent Results*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 369–378.

Nakayama H., *Aspiration Level Approach to Interactive Multi-Objective Programming and Its Applications*, Advances in Multicriteria Analysis, Edited by P.M. Pardalos, Y. Siskos, C. Zopounidis, Kluwer Academic Publishers, Dordrecht, 1995, pp. 147–174.

Nakayama H., *Some Remarks on Dualization in Vector Optimization*, Journal of Multi-Criteria Decision Analysis **5**, No. 3 (1996), 218–225.

Nakayama H., Furukawa K., *Satisficing Trade-Off Method with an Application to Multiobjective Structural Design*, Large Scale Systems **8** (1985), 47–57.

Nakayama H., Sawaragi Y., *Satisficing Trade-Off Method for Multiobjective Programming*, Interactive Decision Analysis, Edited by M. Grauer, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 229, Springer-Verlag, 1984, pp. 113–122.

Nakayama H., Tanino T., Sawaragi Y., *An Interactive Optimization Method in Multicriteria Decisionmaking*, IEEE Transactions on Systems, Man, and Cybernetics **10**, No. 3 (1980), 163–169.

Nakayama H., Nomura J., Sawada K., Nakajima R., *An Application of Satisficing Trade-Off Method to a Blending Problem of Industrial Materials*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986, pp. 303–313.

Nakayama H., Kaneshige K., Takemoto S., Watada Y., *An Application of a Multi-Objective Programming Technique to Construction Accuracy Control of Cable-Stayed Bridges*, European Journal of Operational Research **87**, No. 3 (1995), 731–738.

Nakayama H., Kaneshige K., Takemoto S., Watada Y., *A Construction Accuracy Control System of Cable Stayed Bridge Using a Multi-Objective Programming Technique*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 501–509.

Narula S.C., Weistroffer H.R., *Algorithms for Multi-Objective Nonlinear Programming Problems: An Overview*, Improving Decision Making in Organisations, Edited by A.G. Lockett, G. Islei, Lecture Notes in Economics and Mathematical Systems 335, Springer-Verlag, Berlin, Heidelberg, 1989a, pp. 434–443.

Narula S.C., Weistroffer H.R., *A Flexible Method for Nonlinear Multicriteria Decisionmaking Problems*, IEEE Transactions on Systems, Man, and Cybernetics **19**, No. 4 (1989b), 883–887.

Narula S.C., Kirilov L., Vassilev V., *An Interactive Algorithm for Solving Multiple Objective Nonlinear Programming Problems*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994a, pp. 119–127.

Narula S.C., Kirilov L., Vassilev V., *Reference Direction Approach for Solving Multiple Objective Nonlinear Programming Problems*, IEEE Transactions on Systems, Man, and Cybernetics **24**, No. 5 (1994b), 804–806.

Ng K.Y., *A Multicriteria Optimization Approach to Aircraft Loading*, Operations Research **40**, No. 6 (1992), 1200–1206.

Nijkamp P., Spronk J., *Interactive Multiple Goal Programming: An Evaluation and Some Results*, Multiple Criteria Decision Making Theory and Applications, Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 177, Springer-Verlag, Berlin, Heidelberg, 1980, pp. 278–293.

Nijkamp P., Rietveld P., Spronk J., *Open Problems in the Operationalization of Multiple Criteria Decision Making*, Systems Analysis Modelling Simulation **5**, No. 4 (1988), 311–322.

Ogryczak W., *Preemptive Reference Point Method*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997a, pp. 156–167.

Ogryczak W., *Reference Distribution – An Interactive Approach to Multiple Homogeneous and Anonymous Criteria*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997b, pp. 156–165.

Ohkubo S., Dissanayake P.B.R., Taniwaki K., *An Approach to Multicriteria Fuzzy Optimization of a Prestressed Concrete Bridge System Considering Cost and Aesthetic Feeling*, Structural Optimization **15**, No. 2 (1998), 132–140.

Ohta H., T. Yamaguchi, *Linear Fractional Goal Programming in Consideration of Fuzzy Solution*, European Journal of Operational Research **92**, No. 1 (1996), 157–165.

Olbrisch M., *The Interactive Reference Point Approach as a Solution Concept for Econometric Decision Models*, X. Symposium on Operations Research, Part 1, Sections 1-5, Edited by M.J. Beckmann, K.-W. Gaede, K. Ritter, H. Schneeweiss, Verlag Anton Hain Meisenheim GmbH, Knigstein/Ts., 1986, pp. 611–619.

Olkucu A., *Conflicting Objectives in Software System Design*, Multiple Criteria Decision Making and Risk Analysis Using Microcomputers, Edited by B. Karpak, S. Zionts, Springer-Verlag, Berlin, Heidelberg, 1989, pp. 357–394.

Olson D.L., *Review of Empirical Studies in Multiobjective Mathematical Programming: Subject Reflection of Nonlinear Utility and Learning*, Decision Sciences **23**, No. 1 (1992), 1–20.

Olson D.L., *Tchebycheff Norms in Multi-Objective Linear Programming*, Mathematical and Computer Modelling **17**, No. 1 (1993), 113–124.

Oppenheimer K.R., *A Proxy Approach to Multi-Attribute Decision Making*, Management Science **24**, No. 6 (1978), 675–689.

Osman M.S.A., Ragab A.M., *A Unified Approach for Characterizing the Noninferior Solutions of Multiobjective Nonlinear Programming Problems*, X. Symposium on Operations Research, Part 1, Sections 1-5, Edited by M.J. Beckmann, K.-W. Gaede, K. Ritter, H. Schneeweiss, Verlag Anton Hain Meisenheim GmbH, Knigstein/Ts., 1986a, pp. 133–141.

Osman M.S.A., Ragab A.M., *An Algorithm for Solving Multiobjective Nonlinear Programming Problems*, X. Symposium on Operations Research, Part 1, Sections 1-5, Edited by M.J. Beckmann, K.-W. Gaede, K. Ritter, H. Schneeweiss, Verlag Anton Hain Meisenheim GmbH, Knigstein/Ts., 1986b, pp. 143–147.

Osyczka A., *Multicriterion Optimization in Engineering with FORTRAN Programs*, Ellis Horwood Limited, 1984.

Osyczka A., *Multicriterion Decision Making with Min-Max Approach*, Optimization Methods in Structural Design, Edited by H. Eschenauer, N. Olhoff, Euromech-Colloquium 164, Wissenschaftsverlag, 1989a.

Osyczka A., *Computer Aided Multicriterion Optimization System (CAMOS)*, Discretization Methods and Structural Optimization – Procedures and Applications, Edited by H.A. Eschenauer, G. Thierauf, Lecture Notes in Engineering, Springer-Verlag, Berlin, Heidelberg, 1989b, pp. 263–270.

Osyczka A., *Computer Aided Multicriterion Optimization System (CAMOS): Software Package in FORTRAN*, International Software Publishers, 1992.

Osyczka A., Koski J., *Selected Works Related to Multicriterion Optimization Methods for Engineering Design*, Optimization Methods in Structural Design, Edited by H. Eschenauer, N. Olhoff, Euromech-Colloquium 164, Wissenschaftsverlag, 1989.

Osyczka A., Kundu S., *A New Method to Solve Generalized Multicriteria Optimization Problems Using the Simple Genetic Algorithm*, Structural Optimization **10**, No. 2 (1995), 94–99.

Osyczka A., Zajac J., *Multicriteria Optimization of Computationally Expensive Functions and Its Application to Robot Spring Balancing Mechanism Design*, Multicriteria Design Optimization Procedures and Applications, Edited by H. Eschenauer, J. Koski, A. Osyczka, Springer-Verlag, Berlin, Heidelberg, 1990, pp. 168–183.

Ozernoy V.M., *Choosing the "Best" Multiple Criteria Decision-Making Method*, INFOR **30**, No. 2 (1992a), 159–171.

Ozernoy V.M., *Developing an Expert System Prototype for Intelligent Decision Support*, Multiple Criteria Decision Making: Proceedings of the Ninth International Conference: Theory and Applications in Business, Industry, and Government, Edited by A. Goicoechea, L. Duckstein, S. Zionts, Springer-Verlag, New York, 1992b, pp. 325–333.

Pal B.B., Basu I., *A Goal Programming Method for Solving Fractional Programming Problems via Dynamic Programming*, Optimization **35**, No. 2 (1995), 145–157.

Papageorgiou N.S., *Pareto Efficiency in Locally Convex Spaces II: Stability Results*, Numerical Functional Analysis and Optimization **8**, No. 1 (1985), 117–136.

Pareto V., *Cours d'Economie Politique*, Libraire Droz, Genève, 1964 (the first edition in 1896).

Pareto V., *Manual of Political Economy*, The MacMillan Press Ltd, 1971 (the original edition in French in 1927).

Pascoletti A., Serafini P., *Scalarizing Vector Optimization Problems*, Journal of Optimization Theory and Applications **42**, No. 4 (1984), 499–524.

Passy U., Levanon Y., *Analysis of Multiobjective Decision Problems by the Indifference Band Approach*, Journal of Optimization Theory and Applications **43**, No. 2 (1984), 205–235.

Peterson E.S., *Applications of Multiobjective Decision Making in the Mining Industry*, Papers on Nonlinear Analysis and Mathematical Systems Theory, Edited by J. Szép, P. Tallós, Department of Mathematics, Karl Marx University of Economics, Budapest, 1984, pp. 37–65.

Plastria F., Carrizosa E., *Geometrical Characterization of Weakly Efficient Points*, Journal of Optimization Theory and Applications **90**, No. 1 (1996), 217–223.

Polak E., Payne A.N., *On Multicriteria Optimization*, Directions in Large-Scale Systems, Many-Person Optimization and Decentralized Control, Edited by Y.C. Ho, S.K. Mitter, Plenum Press, New York, 1976, pp. 77–94.

Powell J.G., Premachandra I.M., *Accommodating Diverse Institutional Investment Objectives and Constraints Using Non-Linear Goal Programming*, European Journal of Operational Research **105**, No. 3 (1998), 447–456.

Preda V., *On Efficiency and Duality for Multiobjective Programs*, Journal of Mathematical Analysis and Applications **166**, No. 2 (1992), 365–377.

Preda V., *Optimality Conditions and Duality in Multiple Objective Programming Involving Semilocally Convex and Related Functions*, Optimization **36**, No. 3 (1996), 219–230.

Preda V., Stancu-Minasian I.M., *Generalized Mond-Weir Duality for Multiobjective Nonsmooth Programming*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 55–63.

Rall L.B., *Automatic Differentiation: Techniques and Applications*, Lecture Notes in Computer Science 120, Springer-Verlag, Berlin, Heidelberg, 1981.

Ramesh R., Karwan M.H., Zionts S., *Theory of Convex Cones in Multicriteria Decision Making*, Annals of Operations Research **16**, No. 1–4 (1988), 131–147.

Ramesh R., Karwan M.H., Zionts S., *Interactive Multicriteria Linear Programming: An Extension of the Method of Zionts and Wallenius*, Naval Research Logistics **36**, No. 3 (1989a), 321–335.

Ramesh R., Karwan M.H., Zionts S., *Preference Structure Representation Using Convex Cones in Multicriteria Integer Programming*, Management Science **35**, No. 9 (1989b), 1092–1105.

Rao S.S., *Optimization Theory and Applications*, Second Edition, Wiley Eastern Limited, 1984.

Rao S.S., *Game Theory Approach for Multiobjective Structural Optimization*, Computers and Structures **25**, No. 1 (1987), 119–127.

Rarig H.M., Haimes Y.Y., *Risk/Dispersion Index Method*, IEEE Transactions on Systems, Man, and Cybernetics **13**, No. 3 (1983), 317–328.

Reeves G.R., Gonzalez J.J., *A Comparison of Two Interactive MCDM Procedures*, European Journal of Operational Research **41**, No. 2 (1989), 203–209.

Reeves G.R., Reid R.C., *Minimum Values over the Efficient Set in Multiple Objective Decision Making*, European Journal of Operational Research **36**, No. 3 (1988), 334–338.

ReVelle C., *Equalizing Superpower Force Disparities with Optimized Arms Control Choices: A Multi-Objective Approach*, European Journal of Operational Research **33**, No. 1 (1988), 46–53.

Rietveld P., *Multiple Objective Decision Methods and Regional Planning*, North-Holland Publishing Company, 1980.

Ringuest J.L., *Multiobjective Optimization: Behavioral and Computational Considerations*, Kluwer Academic Publisher, 1992.

Rogowski T., Sobczyk J., Wierzbicki A., *IAC-DIDAS-L: A Dynamic Interactive Decision Analysis and Support System for Multicriteria Analysis of Linear and Dynamic Linear Models of Professional Microcomputers*, Theory, Software and Testing Examples for Decision Support Systems, Edited by A. Lewandowski, A. Wierzbicki, Working Paper WP-87-26, IIASA, Laxenburg, 1987, pp. 106–124.

Romero C., *Handbook of Critical Issues in Goal Programming*, Pergamon Press, 1991.

Romero C., *Goal Programming and Multiple Criteria Decision Making: Some Reflections*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 192–198.

Rosenthal R.E., *Goal Programming – A Critique*, New Zealand Operational Research **11**, No. 1 (1983), 1–7.

Rosenthal R.E., *Principles of Multiobjective Optimization*, Decision Sciences **16**, No. 2 (1985), 133–152.

Rosinger E.E., *Interactive Algorithm for Multiobjective Optimization*, Journal of Optimization Theory and Applications **35**, No. 3 (1981), 339–365; *Errata Corrige*, Journal of Optimization Theory and Applications **38**, No. 1 (1982), 147–148.

Rosinger E.E., *Aids for Decision Making with Conflicting Objectives*, Mathematics of Multi Objective Optimization, Edited by P. Serafini, Springer-Verlag, Wien, New York, 1985, pp. 275–315.

Rothermel M.A., Schilling D.A., *Conjoint Measurement in Multiple Objective Decision Making: A New Approach*, European Journal of Operational Research **23**, No. 3 (1986), 310–319.

Roy A., Mackin P., *Multicriteria Optimization (Linear and Nonlinear) Using Proxy Value Functions*, Multiple Criteria Decision Support, Edited by P. Korhonen, A. Lewandowski, J. Wallenius, Lecture Notes in Economics and Mathematical Systems 356, Springer-Verlag, 1991, pp. 128–134.

Roy A., Wallenius J., *Nonlinear Multiobjective Optimization: An Algorithm and Some Theory*, Mathematical Programming **55**, No. 2 (1992), 235–249.

Roy B., *The Outranking Approach and the Foundations of ELECTRE Methods*, Readings in Multiple Criteria Decision Aid, Edited by C.A. Bana e Costa, Springer-Verlag, Berlin, Heidelberg, 1990, pp. 155–183.

Roy B., Mousseau V., *A Theoretical Framework for Analysing the Notion of Relative Importance of Criteria*, Journal of Multi-Criteria Decision Analysis **5**, No. 2 (1996), 145–159.

Ruíz-Canales P., Rufián-Lizana A., *A Characterization of Weakly Efficient Points*, Mathematical Programming **68**, No. 2 (1995), 205–212.

Saber H.M., Ravindran A., *A Partitioning Gradient Based (PGB) Algorithm for Solving Nonlinear Goal Programming Problems*, Computers & Operations Research **23**, No. 2 (1996), 141–152.

Sadagopan S., Ravindran A., *Interactive Algorithms for Multiple Criteria Nonlinear Programming Problems*, European Journal of Operational Research **25**, No. 2 (1986), 247–257.

Sadek I.S., Bruch J.C. Jr., Sloss J.M., Adali S., *Structural Control of a Variable Cross-Section Beam by Distributed Forces*, Mechanics of Structures and Machines **16**, No. 3 (1988–89), 313–333.

Sainfort F.C., Gustafson D.H., Bosworth K., Hawkins R.P., *Decision Support Systems Effectiveness: Conceptual Framework and Empirical Evaluation*, Organizational Behavior and Human Decision Processes **45** (1990), 232–252.

Sakawa M., *Interactive Multiobjective Decision Making by the Sequential Proxy Optimization Technique: SPOT*, European Journal of Operational Research **9**, No. 4 (1982), 386–396.

Sakawa M., Mori N., *Interactive Multiobjective Decisionmaking for Nonconvex Problems Based on the Weighted Tchebycheff Norm*, Large Scale Systems **5**, No. 1 (1983), 69–82.

Sakawa M., Mori N., *Interactive Multiobjective Decisionmaking for Nonconvex Problems Based on the Penalty Scalarizing Functions*, European Journal of Operational Research **17**, No. 3 (1984), 320–330.

Sakawa M., Seo F., *Interactive Multiobjective Decisionmaking for Large-Scale Systems and Its Application to Environmental Systems*, IEEE Transactions on Systems, Man, and Cybernetics **10**, No. 12 (1980), 796–806.

Sakawa M., Seo F., *Interactive Multiobjective Decision Making by the Sequential Proxy Optimization Technique (SPOT) and Its Application to Environmental Systems*, Control Science and Technology for the Progress of Society, Edited by H. Akashi, IFAC, vol. 2, Pergamon Press, 1982a, pp. 1527–1532.

Sakawa M., Seo F., *Interactive Multiobjective Decision Making in Environmental Systems Using Sequential Proxy Optimization Techniques (SPOT)*, Automatica **18**, No. 2 (1982b), 155–165.

Sakawa M., Yano H., *Interactive Multiobjective Reliability Design of a Standby System by the Fuzzy Sequential Proxy Optimization Technique (FSPOT)*, International Journal of Systems Science **16**, No. 2 (1985), 177–195.

Sakawa M., Yano H., *Trade-Off Rates in the Hyperplane Method for Multiobjective Optimization Problems*, European Journal of Operational Research **44**, No. 1 (1990), 105–118.

Sakawa M., Yano H., *Generalized Hyperplane Methods for Characterizing Λ-Extreme Points and Trade-Off Rates for Multiobjective Optimization Problems*, European Journal of Operational Research **57**, No. 3 (1992), 368–380.

Sankaran S., *Multiple Objective Decision Making Approach to Cell Formation: A Goal Programming Model*, Mathematical and Computer Modelling **13**, No. 9 (1990), 71–81.

Sarma G.V., Merouani H.F., *Some Difficulties in Applying Interactive Compromise Programming Illustrated by a Recent Method and a Case Study*, Journal of the Operational Research Society **46**, No. 1 (1995), 9–19.

Sawaragi Y., Nakayama H., Tanino T., *Theory of Multiobjective Optimization*, Academic Press, Inc., Orlando, Florida, 1985.

Schilling D.A., ReVelle C., Cohon J., *An Approach to the Display and Analysis of Multiobjective Problems*, Socio-Economic Planning Sciences **17**, No. 2 (1983), 57–63.

Schniederjans M.J., *Goal Programming: Methodology and Applications*, Kluwer Academic Publishers, Boston, Dordrecht, London, 1995a.

Schniederjans M.J., *The Life Cycle of Goal Programming Research as Recorded in Journal Articles*, Operations Research **43**, No. 3 (1995b), 551–557.

Schniederjans M.J., Hoffman J., *Multinational Acquisition Analysis: A Zero-One Goal Programming Model*, European Journal of Operational Research **62**, No. 2 (1992), 175–185.

Sherali H.D., *Equivalent Weights for Lexicographic Multi-Objective Programs: Characterizations and Computations*, European Journal of Operational Research **11**, No. 4 (1982), 367–379.

Sherali H.D., Soyster A.L., *Preemptive and Nonpreemptive Multi-Objective Programming: Relationships and Counterexamples*, Journal of Optimization Theory and Applications **39**, No. 2 (1983), 173–186.

Shimizu K., Ishizuka Y., Bard J.F., *Nondifferentiable and Two-Level Mathematical Programming*, Kluwer Academic Publishers, 1997.

Shin W.S., Ravindran A., *Interactive Multiple Objective Optimization: Survey I – Continuous Case*, Computers & Operations Research **18**, No. 1 (1991), 97–114.

Shin W.S., Ravindran A., *A Comparative Study of Interactive Tradeoff Cutting Plane Methods for MOMP*, European Journal of Operational Research **56**, No. 3 (1992), 380–393.

Shor N.Z., *Minimization Methods for Non-Differentiable Functions*, Springer-Verlag, Berlin, 1985.

Silverman J., Steuer R.E., Whisman A.W., *Computer Graphics at the Multicriterion Computer/User Interface*, Decision Making with Multiple Objectives, Edited by Y.Y. Haimes, V. Chankong, Lecture Notes in Economics and Mathematical Systems 242, Springer-Verlag, Berlin, Heidelberg, 1985, pp. 201–213.

Silverman J., Steuer R.E., Whisman A.W., *A Multi-Period, Multiple Criteria Optimization System for Manpower Planning*, European Journal of Operational Research **34**, No. 2 (1988), 160–170.

Simon C.P., *Scalar and Vector Maximization: Calculus Techniques with Economics Applications*, Studies in Mathematical Economics, Edited by S. Reiter, MAA Studies in Mathematics 25, The Mathematical Association of America, 1986, pp. 62–159.

Singh C., Bhatia D., Rueda N., *Duality in Nonlinear Multiobjective Programming Using Augmented Lagrangian Functions*, Journal of Optimization Theory and Applications **88**, No. 3 (1996), 659–670.

Sinha S.B., Sastry S.V.C., Misra R.P., *Management of a Community Storage System: A Goal Programming Model*, European Journal of Operational Research **34**, No. 1 (1988), 92–98.

Skulimowski A.M.J., *Classification and Properties of Dominating Points in Vector Optimization*, Methods of Operations Research 58, Edited by P. Kleinschmidt, F.J. Radermacher, W. Schweitzer, H. Wildermann, Athenum Verlag, Frankfurt am Main, 1989, pp. 99–112.

Skulimowski A.M.J., *Applicability of Ideal Points in Multicriteria Decision Making*, Multiple Criteria Decision Making: Proceedings of the Ninth International Conference: Theory and Applications in Business, Industry, and Government, Edited by A. Goicoechea, L. Duckstein, S. Zionts, Springer-Verlag, New York, 1992, pp. 387–400.

Skulimowski A.M.J., *Decision Support Systems Based on Reference Sets*, Wydawnictwa AGH, Kraków, 1996.

Slowinski R., *Interactive Multiobjective Optimization Based on Ordinal Regression*, Multiobjective Problems of Mathematical Programming, Edited by A. Lewandowski, V. Volkovich, Lecture Notes in Economics and Mathematical Systems 351, Springer-Verlag, 1991, pp. 93–100.

Slowinski R., Teghem J. (Eds.), *Stochastic versus Fuzzy Approaches to Multiobjective Mathematical Programming under Uncertainty*, Kluwer Academic Publishers, Dordrecht, 1990.

Sobol I.M., *An Efficient Approach to Multicriteria Optimum Design Problems*, Surveys on Mathematics for Industry **1**, No. 4 (1992), 259–281.

Sobol' I.M., Levitan Y.L., *Error Estimates for the Crude Approximation of the Trade-off Curve*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 83–92.

Soland R.M., *Multicriteria Optimization: A General Characterization of Efficient Solutions*, Decision Sciences **10**, No. 1 (1979), 26–38.

Soloveychik D., *The Multiple Criteria Decision-Making Process Using the Methods of Multivariate Statistical Analysis*, Fifth International Conference on Collective Phenomena, Edited by J.L. Lebowitz, The New York Academy of Sciences, New York, 1983, pp. 205–212.

Song A., Cheng W.-M., *A Method for Multihuman and Multi-Criteria Decision Making*, Systems Analysis and Simulation 1988 I: Theory and Foundations, Edited by A. Sydow, S.G. Tzafestas, R. Vichnevetsky, Akademie-Verlag, Berlin, 1988, pp. 213–216.

Sounderpandian J., *Value Functions When Decision Criteria Are Not Totally Substitutable*, Operations Research **39**, No. 4 (1991), 592–600.

Soyibo A., *Goal Programming Methods and Applications: A Survey*, Journal of Information and Optimization Sciences **6**, No. 3 (1985), 247–264.

Spronk J., *Interactive Multifactorial Planning: State of the Art*, Readings in Multiple Criteria Decision Aid, Edited by C.A. Bana e Costa, Springer-Verlag, Berlin, Heidelberg, 1990, pp. 512–534.

Stadler W., *A Survey of Multicriteria Optimization or the Vector Maximum Problem, Part I: 1776-1960*, Journal of Optimization Theory and Applications **29**, No. 1 (1979), 1–52.

Stadler W. (Ed.), *Multicriteria Optimization in Engineering and in the Sciences*, Plenum Press, New York, 1988a.

Stadler W., *Fundamentals of Multicriteria Optimization*, Multicriteria Optimization in Engineering and in the Sciences, Edited by W. Stadler, Plenum Press, New York, 1988b, pp. 1–25.

Staib T., *On Necessary and Sufficient Optimality Conditions for Multicriterial Optimization Problems*, ZOR – Methods and Models of Operations Research **35**, No. 3 (1991), 231–248.

Stam A., Lee Y.-R., Yu P.L., *Value Functions and Preference Structures*, Mathematics of Multiobjective Optimization, Edited by P. Serafini, CISM Courses and Lectures 289, 1985, pp. 1–22.

Stam A., Kuula M., Cesar H., *Transboundary Air Pollution in Europe: An Interactive Multicriteria Tradeoff Analysis*, European Journal of Operational Research **56**, No. 2 (1992), 263–277.

Stancu-Minasian I.M., *Stochastic Programming with Multiple Objective Functions*, Mathematics and Its Applications (East European Series), D. Reidel Publishing Company; Editura Academiei, 1984.

Statnikov R.B., Matusov J., *Use of P_r-Nets for the Approximation of the Edgeworth-Pareto Set in Multicriteria Optimization*, Journal of Optimization Theory and Applications **91**, No. 3 (1996), 543–560.

Sterna-Karwat A., *Continuous Dependence of Solutions on a Parameter in a Scalarization Method*, Journal of Optimization Theory and Applications **55**, No. 3 (1987), 417–434.

Steuer R.E., *Multiple Criteria Optimization: Theory, Computation, and Applications*, John Wiley & Sons, Inc., 1986.

Steuer R.E., *The Tchebycheff Procedure of Interactive Multiple Objective Programming*, Multiple Criteria Decision Making and Risk Analysis Using Microcomputers, Edited by B. Karpak, S. Zionts, Springer-Verlag, Berlin, Heidelberg, 1989a, pp. 235–249.

Steuer R.E., *Trends in Interactive Multiple Objective Programming*, Methodology and Software for Interactive Decision Support, Edited by A. Lewandowski, I. Stanchev, Lecture Notes in Economics and Mathematical Systems 337, Springer-Verlag, 1989b, pp. 107–119.

Steuer R.E., *Implementing the Tchebycheff Method in a Spreadsheet*, Essays in Decision Making: A Volume in Honour of Stanley Zionts, Edited by M.H. Karwan, J. Spronk, J. Wallenius, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 93–103.

Steuer R.E., Choo E.-U., *An Interactive Weighted Tchebycheff Procedure for Multiple Objective Programming*, Mathematical Programming **26**, No. 3 (1983), 326–344.

Steuer R.E., Gardiner L.R., *Interactive Multiple Objective Programming: Concepts, Current Status, and Future Directions*, Readings in Multiple Criteria Decision Aid, Edited by C.A. Bana e Costa, Springer-Verlag, Berlin, Heidelberg, 1990, pp. 413–444.

Steuer R.E., Gardiner L.R., *On the Computational Testing of Procedures for Interactive Multiple Objective Linear Programming*, Operations Research, Edited by G. Fandel, H. Gehring, Springer-Verlag, Berlin, Heidelberg, 1991, pp. 121–131.

Steuer R.E., Sun M., *The Parameter Space Investigation Method of Multiple Objective Nonlinear Programming: A Computational Investigation*, Operations Research **43**, No. 4 (1995), 641–648.

Steuer R.E., Whisman A.W., *Toward the Consolidation of Interactive Multiple Objective Programming Procedures*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986, pp. 232–241.

Steuer R.E., Silverman J., Whisman A.W., *A Combined Tchebycheff/Aspiration Criterion Vector Interactive Multiobjective Programming Procedure*, Management Science **39**, No. 10 (1993), 1255–1260.

Steuer R.E., Gardiner L.R., Gray J., *A Bibliographic Survey of the Activities and International Nature of Multiple Criteria Decision Making*, Journal of Multi-Criteria Decision Analysis **5**, No. 3 (1996), 195–217.

Stewart T.J., *A Critical Survey on the Status of Multiple Criteria Decision Making Theory and Practice*, OMEGA **20**, No. 5–6 (1992), 569–586.

Stewart T.J., *Convergence and Validation of Interactive Methods in MCDM: Simulation Studies*, Essays in Decision Making: A Volume in Honour of Stanley Zionts, Edited by M.H. Karwan, J. Spronk, J. Wallenius, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 7–18.

Sultan A.M., Templeman A.B., *Generation of Pareto Solutions by Entropy-Based Methods*, Multi-Objective Programming and Goal Programming: Theories and Applications, Edited by M. Tamiz, Lecture Notes in Economics and Mathematical Systems 432, Springer-Verlag, Berlin, Heidelberg, 1996, pp. 164–195.

Sun M., Stam A., Steuer R.E., *Solving Multiple Objective Programming Problems Using Feed-Forward Artificial Neural Networks: The Interactive FFANN Procedure*, Management Science **42**, No. 6 (1996), 835–849.

Sunaga T., Mazeed M.A., Kondo E., *A Penalty Function Formulation for Interactive Multiobjective Programming Problems*, System Modelling and Optimization, Edited by M. Iri, K. Yajima, Lecture Notes in Control and Information Sciences 113, Springer-Verlag, 1988, pp. 221–230.

Szidarovszky F., Szenteleki K., *A Multiobjective Optimization Model for Wine Production*, Applied Mathematics and Computation **22**, No. 2–3, (1987), 255–275.

Tabucanon M.T., *Multiple Criteria Decision Making in Industry*, Elsevier Science Publishers B.V., Amsterdam, 1988.

Tamiz M., Jones D.F., *Algorithmic Improvements to the Method of Martel and Aouni*, Journal of the Operational Research Society **46**, No. 2 (1995), 254–257.

Tamiz M., Jones D.F., *An Overview of Current Solution Methods and Modelling Practices in Goal Programming*, Multi-Objective Programming and Goal Programming: Theories and Applications, Edited by M. Tamiz, Lecture Notes in Economics and Mathematical Systems 432, Springer-Verlag, Berlin, Heidelberg, 1996, pp. 198–211.

Tamiz M., Jones D.F., *Interactive Frameworks for Investigation of Goal Programming Models: Theory and Practice*, Journal of Multi-Criteria Decision Analysis **6**, No. 1 (1997a), 52–60.

Tamiz M., Jones D.F., *A General Interactive Goal Programming Algorithm*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997b, pp. 433–444.

Tamiz M., Hasham R., Jones D.F., Hesni B., Fargher E.K., *A Two Staged Goal Programming Model for Portfolio Selection*, Multi-Objective Programming and Goal Programming: Theories and Applications, Edited by M. Tamiz, Lecture Notes in Economics and Mathematical Systems 432, Springer-Verlag, Berlin, Heidelberg, 1996, pp. 286–299.

Tamura K., Arai S., *On Proper and Improper Efficient Solutions of Optimal Problems with Multicriteria*, Journal of Optimization Theory and Applications **38**, No. 2 (1982), 191–205.

Tan Y.S., Fraser N.M., *The Modified Star Graph and the Petal Diagram: Two New Visual Aids for Discrete Alternative Multicriteria Decision Making*, Journal of Multi-Criteria Decision Analysis **7**, No. 1 (1998), 20–33.

Tanino T., *Sensitivity Analysis in Multiobjective Optimization*, Journal of Optimization Theory and Applications **56**, No. 3 (1988a), 479–499.

Tanino T., *Stability and Sensitivity Analysis in Convex Vector Optimization*, SIAM Journal on Control and Optimization **26**, No. 3 (1988b), 521–536.

Tanino T., *Stability and Sensitivity Analysis in Multiobjective Nonlinear Programming*, Annals of Operations Research **27**, No. 1–4 (1990), 97–114.

Tanino T., Sawaragi Y., *Stability of Nondominated Solutions in Multicriteria Decision-Making*, Journal of Optimization Theory and Applications **30**, No. 2 (1980), 229–253.

Tapia C.G., Murtagh B.A., *The Use of Preference Criteria in Interactive Multiobjective Mathematical Programming*, Asia-Pacific Journal of Operational Research **6**, No. 2 (1989), 131–147.

Tapia C.G., Murtagh B.A., *A Markovian Process in Interactive Multiobjective Decision-Making*, European Journal of Operational Research **57**, No. 3 (1992), 421–428.

Tarvainen K., *On the Implementation of the Interactive Surrogate Worth Trade-Off (ISWT) Method*, Interactive Decision Analysis, Edited by M. Grauer, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 229, Springer-Verlag, Berlin, Heidelberg, 1984, pp. 154–161.

Tarvainen K., *Duality Theory for Preferences in Multiobjective Decisionmaking*, Journal of Optimization Theory and Applications **88**, No. 1 (1996), 237–245.

Tecle A., Duckstein L., *A Procedure for Selecting MCDM Techniques for Forest Resources Management*, Multiple Criteria Decision Making: Proceedings of the Ninth International Conference: Theory and Applications in Business, Industry, and Government, Edited by A. Goicoechea, L. Duckstein, S. Zionts, Springer-Verlag, New York, 1992, pp. 19–32.

Teghem J. Jr., Delhaye C., Kunsch P.L., *An Interactive Decision Support System (IDSS) for Multicriteria Decision Aid*, Mathematical and Computer Modelling **12**, No. 10–11 (1989), 1311–1320.

Tell R., Wallenius J., *A Survey of Multiple-Criteria Decision Methods and Applications: Utility Theory and Mathematical Programming*, The Finnish Journal of Business Economics **1**, No. 1 (1979), 3–22.

Tenhuisen M.L., Wiecek M.M., *On the Structure of the Non-Dominated Set for Bicriteria Programmes*, Journal of Multi-Criteria Decision Analysis **5**, No. 3 (1996), 232–243.

Thach P.T., Konno H., Yokota D., *Dual Approach to Minimization on the Set of Pareto-Optimal Solutions*, Journal of Optimization Theory and Applications **88**, No. 3 (1996), 689–707.

Thore S., Nagurney A., Pan J., *Generalized Goal Programming and Variational Inequalities*, Operations Research Letters **12**, No. 4 (1992), 217–226.

Törn A., *A Sampling-Search-Clustering Approach for Solving Scalar (Local, Global) and Vector Optimization Problems*, Theory and Practice of Multiple Criteria Decision Making, Edited by C. Carlsson, Y. Kochetkova, North-Holland Publishing Company, 1983, pp. 119–141.

Udink ten Cate A., *On the Determination of the Optimal Temperature for the Growth of an Early Cucumber Crop in a Greenhouse*, Plural Rationality and Interactive Decision Processes, Edited by M. Grauer, M. Thompson, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 248, Springer-Verlag, Berlin, Heidelberg, 1985, pp. 311–318.

Vanderpooten D., *The Use of Preference Information in Multiple Criteria Interactive Procedures*, Improving Decision Making in Organisations, Edited by A.G. Lockett, G. Islei, Lecture Notes in Economics and Mathematical Systems 335, Springer-Verlag, Berlin, Heidelberg, 1989a, pp. 390–399.

Vanderpooten D., *The Interactive Approach in MCDA: A Technical Framework and Some Basic Conceptions*, Mathematical and Computer Modelling **12**, No. 10–11 (1989b), 1213–1220.

Vanderpooten D., *Multiobjective Programming: Basic Concepts and Approaches*, Stochastic versus Fuzzy Approaches to Multiobjective Mathematical Programming under Uncertainty, Edited by R. Slowinski, J. Teghem, Kluwer Academic Publishers, Dordrecht, 1990, pp. 7–22.

Vanderpooten D., *Three Basic Conceptions Underlying Multiple Criteria Interactive Procedures*, Multiple Criteria Decision Making: Proceedings of the Ninth International Conference: Theory and Applications in Business, Industry, and Government, Edited by A. Goicoechea, L. Duckstein, S. Zionts, Springer-Verlag, New York, 1992, pp. 441–448.

Vanderpooten D., Vincke P., *Description and Analysis of Some Representative Interactive Multicriteria Procedures*, Mathematical and Computer Modelling **12**, No. 10–11 (1989), 1221–1238.

Vassilev V., Sgurev V., Atanassov A., Deianov A., Kichovich M., Kirilov L., Djambov V., Vachkov G., *Software Product for Multiobjective Nonlinear Programming: MONP-16*, Version 1.1, General Description; User's Guide, Software Products and Systems Corporation, Institute of Industrial Cybernetics and Robotics, Sofia, 1990.

Verkama M., Heiskanen P., *Comment on a Decision Support Approach for Negotiation: Software vs. Methodology*, European Journal of Operational Research **96**, No. 1 (1996), 202–204.

Vessey I., *Cognitive Fit: A Theory-Based Analysis of the Graphs versus Tables Literature*, Decision Sciences **22** (1991), 219–240.

Vetschera R., *Feedback-Oriented Group Decision Support in a Reference Point Framework*, Multiobjective Problems of Mathematical Programming, Edited by A. Lewandowski, V. Volkovich, Lecture Notes in Economics and Mathematical Systems 351, Springer-Verlag, 1991a, pp. 309–314.

Vetschera R., *A Note on Scalarizing Functions under Changing Sets of Criteria*, European Journal of Operational Research **52**, No. 1 (1991b), 113–118.

Vetschera R., *A Preference-Preserving Projection Technique for MCDM*, European Journal of Operational Research **61**, No. 1–2 (1992), 195–203.

Vincke P., *Multicriteria Decision-Aid*, John Wiley & Sons, Inc., Chichester, 1992.

Wallenius J., *Comparative Evaluation of Some Interactive Approaches to Multicriterion Optimization*, Management Science **21**, No. 12 (1975), 1387–1396.

Wallenius J., Zionts S., *A Research Project on Multicriteria Decision Making*, Conflicting Objectives in Decision, Edited by D.E. Bell, R. Keeney, H. Raiffa, John Wiley & Sons, Inc., New York, 1977.

Wan Y.-H., *On Local Pareto Optima*, Journal of Mathematical Economics **2**, No. 1 (1975), 35–42.

Wang S., *Lagrange Conditions in Nonsmooth and Multiobjective Mathematical Programming*, Mathematics in Economics **1** (1984), 183–193.

Wang S., *Algorithms for Multiobjective and Nonsmooth Optimization*, Methods of Operations Research 58, Edited by P. Kleinschmidt, F.J. Radermacher, W. Schweitzer, H. Wildermann, Athenum Verlag, Frankfurt am Main, 1989, pp. 131–142.

Wang S., *Second-Order Necessary and Sufficient Conditions in Multiobjective Programming*, Numerical Functional Analysis and Application **12**, No. 1–2 (1991), 237–252.

Wang S., *An Interactive Method for Multicriteria Decision Making*, Optimization: Techniques and Applications, vol. 1, Proceedings of the International Conference (ICOTA), Edited by K.H. Phua, C.M. Wang, W.Y. Yeong, T.Y. Leong, H.T. Loh, K.C. Tan, F.S. Chou, World Scientific Publishing Co., Inc., 1992, pp. 307–316.

Wang S., Li Z., *Scalarization and Lagrange Duality in Multiobjective Optimization*, Optimization **26**, No. 3–4 (1992), 315–324.

Wang S., Zhou X., *Multiobjective Optimization in Spillway Profile*, Approximation, Optimization and Computing: Theory and Applications, Edited by A.G. Law, C.L. Wang, Elsevier Science Publishers B.V., 1990, pp. 309–312.

Warburton A., *Quasiconcave Vector Maximization: Connectedness of the Sets of Pareto-Optimal and Weak Pareto-Optimal Alternatives*, Journal of Optimization Theory and Applications **40**, No. 4 (1983), 537–557.

Weber M., *Decision Making with Incomplete Information*, European Journal of Operational Research **28**, No. 1 (1987), 44–57.

Weck M., Förtsch F., *Application of Multicriteria Optimization to Structural Systems*, System Modelling and Optimization, Edited by M. Iri, K. Yajima, Lecture Notes in Control and Information Sciences 113, Springer-Verlag, 1988, pp. 471–483.

Weidner P., *On the Characterization of Efficient Points by Means of Monotone Functionals*, Optimization **19**, No. 1 (1988), 53–69.

Weidner P., *Complete Efficiency and Interdependencies Between Objective Functions in Vector Optimization*, ZOR – Methods and Models of Operations Research **34**, No. 2 (1990), 91–115.

Weir T., *Proper Efficiency and Duality for Vector Valued Optimization Problems*, Journal of the Australian Mathematical Society, Series A **43**, No. 1 (1987), 21–34.

Weir T., *On Efficiency, Proper Efficiency and Duality in Multiobjective Programming*, Asia-Pacific Journal of Operational Research **7**, No. 1 (1990), 46–54.

Weistroffer H.R., *Multiple Criteria Decision Making with Interactive Over-Achievement Programming*, Operations Research Letters **1**, No. 6 (1982), 241–245.

Weistroffer H.R., *An Interactive Goal-Programming Method for Non-Linear Multiple-Criteria Decision-Making Problems*, Computers & Operations Research **10**, No. 4 (1983), 311–320.

Weistroffer H.R., *A Combined Over- and Under-Achievement Programming Approach to Multiple Objective Decision-Making*, Large Scale Systems **7**, No. 1 (1984), 47–58.

Weistroffer H.R., *Careful Usage of Pessimistic Values is Needed in Multiple Objective Optimization*, Operations Research Letters **4**, No. 1 (1985), 23–25.

Weistroffer H.R., *A Flexible Model for Multi-Objective Optimization*, Recent Advances and Historical Developments of Vector Optimization, Edited by J. Jahn, W. Krabs, Lecture Notes in Economics and Mathematical Systems 294, Springer-Verlag, Berlin, Heidelberg, 1987, pp. 311–316.

Weistroffer H.R., Narula S., *The Current State of Nonlinear Multiple Criteria Decision Making*, Operations Research, Edited by G. Fandel, H. Gehring, Springer-Verlag, Berlin, Heidelberg, 1991, pp. 109–119.

Weistroffer H.R., Narula S.C., *The State of Multiple Criteria Decision Support Software*, Annals of Operations Research **72** (1997), 299–313.

Wendell R.E., Lee D.N., *Efficiency in Multiple Objective Optimization Problems*, Mathematical Programming **12**, No. 3 (1977), 406–414.

White D.J., *Concepts of Proper Efficiency*, European Journal of Operational Research **13**, No. 2 (1983a), 180–188.

White D.J., *A Selection of Multi-Objective Interactive Programming Methods*, Multi-Objective Decision Making, Edited by S. French, R. Hartley, L.C. Thomas, D.J. White, Academic Press, 1983b, pp. 99–126.

White D.J., *A Bibliography on the Applications of Mathematical Programming Multiple-Objective Methods*, Journal of the Operational Research Society **41**, No. 8 (1990), 669–691.

Wierzbicki A.P., *Basic Properties of Scalarization Functionals for Multiobjective Optimization*, Mathematische Operationsforschung und Statistik, Ser. Optimization **8**, No. 1 (1977), 55–60.

Wierzbicki A.P., *A Methodological Guide to Multiobjective Optimization*, Optimization Techniques, Part 1, Edited by K. Iracki, K. Malanowski, S. Walukiewicz, Lecture Notes in Control and Information Sciences 22, Springer-Verlag, Berlin, Heidelberg, 1980a, pp. 99–123.

Wierzbicki A.P., *The Use of Reference Objectives in Multiobjective Optimization*, Multiple Criteria Decision Making Theory and Applications, Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 177, Springer-Verlag, Berlin, Heidelberg, 1980b, pp. 468–486.

Wierzbicki A.P., *A Mathematical Basis for Satisficing Decision Making*, Organizations: Multiple Agents with Multiple Criteria, Edited by N.N. Morse, Lecture Notes in Economics and Mathematical Systems 190, Springer-Verlag, Berlin, Heidelberg, 1981, pp. 465–485.

Wierzbicki A.P., *A Mathematical Basis for Satisficing Decision Making*, Mathematical Modelling **3**, No. 25 (1982), 391–405.

Wierzbicki A.P., *On the Completeness and Constructiveness of Parametric Characterizations to Vector Optimization Problems*, OR Spektrum **8**, No. 2 (1986a), 73–87.

Wierzbicki A.P., *A Methodological Approach to Comparing Parametric Characterizations of Efficient Solutions*, Large-Scale Modelling and Interactive Decision Analysis, Edited by G. Fandel, M. Grauer, A. Kurzhanski, A.P. Wierzbicki, Lecture Notes in Economics and Mathematical Systems 273, Springer-Verlag, 1986b, pp. 27–45.

Wierzbicki A.P., *On the Role of Intuition in Decision Making and Some Ways of Multicriteria Aid of Intuition*, Journal of Multi-Criteria Decision Analysis **6**, No. 2 (1997a), 65–76.

Wierzbicki A.P., *Convergence of Interactive Procedures of Multiobjective Optmization and Decision Support*, Essays in Decision Making: A Volume in Honour of Stanley Zionts, Edited by M.H. Karwan, J. Spronk, J. Wallenius, Springer-Verlag, Berlin, Heidelberg, 1997b, pp. 19–47.

Wierzbicki A.P., Granat J., *Multi-Objective Modeling for Engineering Application in Decision Support*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 529–540.

Winkels H.-M., Meika M., *An Integration of Efficiency Projections into the Geoffrion Approach for Multiobjective Linear Programming*, European Journal of Operational Research **16**, No. 1 (1984), 113–127.

Wood E., Greis N.P., Steuer R.E., *Linear and Nonlinear Applications of the Tchebycheff Metric to the Multi Criteria Water Allocation Problem*, Environmental Systems Analysis and Management, Edited by S. Rinaldi, North-Holland Publishing Company, 1982, pp. 363–376.

Yang J.-B., Sen P., *Preference Modelling by Estimating Local Utility Functions for Multiobjective Optimization*, European Journal of Operational Research **95**, No. 1 (1996), 115–138.

Yang J.-B., Chen C., Zhang Z.-J., *The Interactive Step Trade-Off Method (ISTM) for Multiobjective Optimization*, IEEE Transactions on Systems, Man, and Cybernetics **20**, No. 3 (1990), 688–695.

Yano H., Sakawa M., *Trade-Off Rates in the Weighted Tchebycheff Norm Method*, Large Scale Systems **13**, No. 2 (1987), 167–177.

Yano H., Sakawa M., *A Unified Approach for Characterizing Pareto Optimal Solutions of Multiobjective Optimization Problems: The Hyperplane Method*, European Journal of Operational Research **39**, No 1 (1989), 61–70.

Yoon K.P., Hwang C.-L., *Multiple Attribute Decision Making: An Introduction*, Sage Publications, Inc., Thousand Oaks, 1995.

Yoshikawa K., Haruna M., Kobayashi K., *Multi-Objective Multi-Level Systems Approach to Planning of Public Works*, Control Science and Technology for the Progress of Society, Edited by H. Akashi, IFAC, vol. 2, Pergamon Press, 1982, pp. 1545–1551.

Youness E.A., *A Direct Approach for Finding All Efficient Solutions for Multiobjective Programming Problems*, European Journal of Operational Research **81**, No. 2 (1995), 440–443.

Yu P.L., *A Class of Solutions for Group Decision Problems*, Management Science **19**, No. 8 (1973), 936–946.

Yu P.L., *Cone Convexity, Cone Extreme Points, and Nondominated Solutions in Decision Problems with Multiobjectives*, Journal of Optimization Theory and Applications **14**, No. 3 (1974), 319–377.

Yu P.L., *Multiple-Criteria Decision Making Concepts, Techniques, and Extensions*, Plenum Press, New York, 1985.

Yu P.L., *Habitual Domains*, Operations Research **39**, No. 6 (1991), 869–876.

Yu P.L., *Toward Expanding and Enriching Domains of Thinking and Application*, Multiple Criteria Decision Making – Proceedings of the Tenth International Conference: Expand and Enrich the Domains of Thinking and Application, Edited by G.H. Tzeng, H.F. Wand, U.P. Wen, P.L. Yu, Springer-Verlag, New York, 1994, pp. 1–9.

Yu P.-L., *Habitual Domains: Freeing Yourself from the Limits on Your Life*, Highwater Editions, Shawnee Mission, Kansas, 1995.

Yu P.-L., Liu L., *A Foundation of Principles for Expanding Habitual Domains*, Essays in Decision Making: A Volume in Honour of Stanley Zionts, Edited by M.H. Karwan, J. Spronk, J. Wallenius, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 185–200.

Zadeh L., *Optimality and Non-Scalar-Valued Performance Criteria*, IEEE Transactions on Automatic Control **8** (1963), 59–60.

Zangwill W.I., *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Inc., 1969.

Zeleny M., *Compromise Programming*, Multiple Criteria Decision Making, Edited by J.L. Cochrane, M. Zeleny, University of South Carolina Press, Columbia, South Carolina, 1973, pp. 262–301.

Zeleny M., *Linear Multiobjective Programming*, Lecture Notes in Economics and Mathematical Systems 95, Springer-Verlag, Berlin, Heidelberg, 1974.

Zeleny M., *The Theory of Displaced Ideal*, Multiple Criteria Decision Making Kyoto 1975, Edited by M. Zeleny, Lecture Notes in Economics and Mathematical Systems 123, Springer-Verlag, Berlin, Heidelberg, 1976, pp. 153–206.

Zeleny M., *Multiple Criteria Decision Making*, McGraw-Hill, Inc., 1982.

Zeleny M., *Stable Patterns from Decision-Producing Networks: New Interfaces of DSS and MCDM*, MCDM WorldScan **3**, No. 2–3 (1989), 6–7.

Zeleny M., *Towards the Tradeoffs-Free Optimality in MCDM*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 596–601.

Zhou X., Mokhtarian F.S., Zlobec S., *A Simple Constraint Qualification in Convex Programming*, Mathematical Programming **61**, No. 3 (1993), 385–397.

Zionts S., *Methods for Solving Management Problems Involving Multiple Objectives*, Multiple Criteria Decision Making Theory and Application, Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 177, Springer-Verlag, Berlin, Heidelberg, 1980, pp. 540–558.

Zionts S., *Multiple Criteria Mathematical Programming: An Updated Overview and Several Approaches*, Multiple Criteria Decision Making and Risk Analysis Using Microcomputers, Edited by B. Karpak, S. Zionts, Springer-Verlag, Berlin, Heidelberg, 1989, pp. 7–60.

Zionts S., *The State of Multiple Criteria Decision Making: Past, Present, and Future*, Multiple Criteria Decision Making: Proceedings of the Ninth International Conference: Theory and Applications in Business, Industry, and Government, Edited by A. Goicoechea, L. Duckstein, S. Zionts, Springer-Verlag, New York, 1992, pp. 33–43.

Zionts S., *Some Thoughts on MCDM: Myths and Ideas*, Multicriteria Analysis, Edited by J. Clímaco, Springer-Verlag, Berlin, Heidelberg, 1997a, pp. 602–607.

Zionts S., *Decision Making: Some Experiences, Myths and Observations*, Multiple Criteria Decision Making: Proceedings of the Twelfth International Conference, Hagen (Germany), Edited by G. Fandel, T. Gal, Lecture Notes in Economics and Mathematical Systems 448, Springer-Verlag, Berlin, Heidelberg, 1997b, pp. 233–241.

Zionts S., Wallenius J., *An Interactive Programming Method for Solving the Multiple Criteria Problem*, Management Science **22**, No. 6 (1976), 652–663.

Zionts S., Wallenius J., *An Interactive Multiple Objective Linear Programming Method for a Class of Underlying Nonlinear Utility Functions*, Management Science **29**, No. 5 (1983), 519–529.

Zlobec S., *Two Characterizations of Pareto Minima in Convex Multicriteria Optimization*, Aplikace Matematiky **29**, No. 5 (1984), 342–349.

Zubiri J., *Scalarization of Vector Optimization Problems via Generalized Chebyshev Norm*, Mathematical Analysis and Systems Theory V., Edited by J. Szép, P. Tallós, Department of Mathematics, Karl Marx University of Economics, Budapest, 1988, pp. 39–42.

INDEX