ADVANCES IN *SOFT COMPUTING*

Tetsuzo Tanino · Tamaki Tanaka
Masahiro Inuiguchi
Editors

Multi-Objective Programming and Goal-Programming

Theory and Applications

Springer

Multi-Objective Programming and Goal Programming

Advances in Soft Computing

Editor-in-chief
Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw, Poland
E-mail: kacprzyk@ibspan.waw.pl
http://www.springer.de/cgi-bin/search-bock.pl?series=4240

Esko Turunen
Mathematics Behind Fuzzy Logic
1999. ISBN 3-7908-1221-8

Robert Fullér
Introduction to Neuro-Fuzzy Systems
2000. ISBN 3-7908-1256-0

Robert John and Ralph Birkenhead (Eds.)
Soft Computing Techniques and Applications
2000. ISBN 3-7908-1257-9

Mieczysław Kłopotek, Maciej Michalewicz
and Sławomir T. Wierzchoń (Eds.)
Intellligent Information Systems
2000. ISBN 3-7908-1309-5

Peter Sinčák, Ján Vaščák, Vladimír Kvasnička
and Radko Mesiar (Eds.)
The State of the Art in Computational Intelligence
2000. ISBN 3-7908-1322-2

Bernd Reusch and Karl-Heinz Temme (Eds.)
Computational Intelligence in Theory and Practice
2000. ISBN 3-7908-1357-5

Rainer Hampel, Michael Wagenknecht,
Nasredin Chaker (Eds.)
Fuzzy Control
2000. ISBN 3-7908-1327-3

Henrik Larsen, Janusz Kacprzyk,
Sławomir Zadrozny, Troels Andreasen,
Henning Christiansen (Eds.)
Flexible Query Answering Systems
2000. ISBN 3-7908-1347-8

Robert John and Ralph Birkenhead (Eds.)
Developments in Soft Computing
2001. ISBN 3-7908-1361-3

Mieczysław Kłopotek, Maciej Michalewicz
and Sławomir T. Wierzchoń (Eds.)
Intelligent Information Systems 2001
2001. ISBN 3-7908-1407-5

Antonio Di Nola and Giangiacomo Gerla (Eds.)
Lectures on Soft Computing and Fuzzy Logic
2001. ISBN 3-7908-1396-6

Tadeusz Trzaskalik and Jerzy Michnik (Eds.)
Multiple Objective and Goal Programming
2002. ISBN 3-7908-1409-1

James J. Buckley and Esfandiar Eslami
An Introduction to Fuzzy Logic and Fuzzy Sets
2002. ISBN 3-7908-1447-4

Ajith Abraham and Mario Köppen (Eds.)
Hybrid Information Systems
2002. ISBN 3-7908-1480-6

Przemysław Grzegorzewski, Olgierd Hryniewicz,
Maria A. Gil (Eds.)
Soft Methods in Probability, Statistics and Data Analysis
2002. ISBN 3-7908-1526-8

Lech Polkowski
Rough Sets
2002. ISBN 3-7908-1510-1

Mieczysław Kłopotek, Maciej Michalewicz
and Sławomir T. Wierzchoń (Eds.)
Intelligent Information Systems 2002
2002. ISBN 3-7908-1509-8

Andrea Bonarini, Francesco Masulli
and Gabriella Pasi (Eds.)
Soft Computing Applications
2002. ISBN 3-7908-1544-6

Leszek Rutkowski, Janusz Kacprzyk (Eds.)
Neural Networks and Soft Computing
2003. ISBN 3-7908-0005-8

Jürgen Franke, Gholamreza Nakhaeizadeh,
Ingrid Renz (Eds.)
Text Mining
2003. ISBN 3-7908-0041-4

Tetsuzo Tanino
Tamaki Tanaka
Masahiro Inuiguchi

Multi-Objective Programming and Goal Programming

Theory and Applications

With 77 Figures
and 48 Tables

Springer

Professor Tetsuzo Tanino
Professor Masahiro Inuiguchi
Osaka University
Graduate School of Engineering
Dept. of Electronics and Information Systems
2-1 Yamada-Oka, Suita
Osaka 565-0871
Japan

Professor Tamaki Tanaka
Niigata University
Graduate School of Science and Technology
Dept. of Mathematics and Information Science
8050, Ikarashi 2-no-cho
Niigata 950-2181
Japan

Supported by the Commemorative Association for the Japan World Exposition (1970)

Cataloging-in-Publication Data applied for
Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in theInternet at <http://dnb.dd.de>

ISSN 16-15-3871
ISBN 978-3-540-00653-4 ISBN 978-3-540-36510-5 (eBook)
DOI 10.1007/978-3-540-36510-5

http://www.springer.de

Originally published by Springer-Verlag Berlin Heidelberg New York in 2003

Cover design: Erich Kirchner, Heidelberg
Typesetting: Digital data supplied by the authors

Printed on acid-free paper 62/3020Rw-5 4 3 2 1 0

Preface

This volume constitutes the proceedings of the Fifth International Conference on Multi-Objective Programming and Goal Programming: Theory & Applications (MOPGP'02) held in Nara, Japan on June 4–7, 2002. Eighty-two people from 16 countries attended the conference and 78 papers (including 9 plenary talks) were presented.

MOPGP is an international conference within which researchers and practitioners can meet and learn from each other about the recent development in multi-objective programming and goal programming. The participants are from different disciplines such as Optimization, Operations Research, Mathematical Programming and Multi-Criteria Decision Aid, whose common interest is in multi-objective analysis.

The first MOPGP Conference was held at Portsmouth, United Kingdom, in 1994. The subsequent conferenes were held at Torremolinos, Spain in 1996, at Quebec City, Canada in 1998, and at Katowice, Poland in 2000. The fifth conference was held at Nara, which was the capital of Japan for more than seventy years in the eighth century. During this Nara period the basis of Japanese society, or culture established itself. Nara is a beautiful place and has a number of historic monuments in the World Heritage List.

The members of the International Committee of MOPGP'02 were Dylan Jones, Pekka Korhonen, Carlos Romero, Ralph Steuer and Mehrdad Tamiz. The Local Committee in Japan consisted of Masahiro Inuiguchi (Osaka University), Hiroataka Nakayama (Konan University), Eiji Takeda (Osaka Unviersity), Hiroyuki Tamura (Osaka University), Tamaki Tanaka (Niigata Unviersity) – co-chair, Tetsuzo Tanino (Osaka University) – co-chair, and Kiichiro Tsuji (osaka University). We would like to thank the secretaries, Keiji Tatsumi (Osaka Unviersity), Masayo Tsurumi (Tokyo University of Science), Syuuji Yamada (Toyama College) and Ye-Boon Yun (Kagawa University) for their earnest work.

We highly appreciate the financial support that the Commemorative Association for the Japan World Exposition (1970) gave us. We would also like to thank the following organizations which have made helpful supports and endorsements for MOPGP'02: The Pacific Optimization Research Activity Group (POP), the Institute of Systems, Control and Information Engineers (ISCIE) and Japan Society for Fuzzy Theory and Systems (SOFT). We are grateful, last but not least, to Nara Convention Bureau for several supports. Particulary, without the devoteful help by Mrs. Keiko Nakamura and Mr. Shigekazu Kuribayashi, this conference would not had been possible.

This volume consists of 61 papers. Thanks to the efforts made by the referees, readers will enjoy turning the pages.

Osaka and Niigata,
December, 2002

Tetsuzo Tanino
Tamaki Tanaka
Masahiro Inuiguchi

Contents

PART I:

Invited Papers

Multiple Objective Combinatorial Optimization – A Tutorial

Matthias Ehrgott[1] and Xavier Gandibleux[2]

[1] Department of Engineering Science, University of Auckland, Private Bag 92019, Auckland, New Zealand, email m.ehrgott@auckland.ac.nz
[2] LAMIH – UMR CNRS 8530, Université de Valenciennes, Campus "Le Mont Houy", F-59313 Valenciennes cedex 9, France, email xavier.gandibleux@univ-valenciennes.fr

Abstract. In this paper we take the reader on a very brief guided tour of multiobjective combinatorial optimization (MOCO). We point out the increasing importance of consideration of multiple objectives in real world applications of combinatorial optimization, survey the problem context and the main characteristics of (MOCO) problems. Our main stops on the tour are for overviews of exact and heuristic solution methods for MOCO. We conclude the presentation with an outlook on promising destinations for future expeditions into the field.

1 Importance in Practice

The importance of multiobjective combinatorial optimization for the solution of real world problems has been recognized in the last few years. We present a number of examples. Trip organization (for tourism purposes) involves minimizing transport, activity, and lodging cost while at the same time maximizing attractivity of activities and lodging. This problem has been formulated as a preference-based multicriteria TSP and heuristic methods have been applied for its solution [39]. In airline crew scheduling the classical objective is to minimize cost. However, minimal cost crew schedules might be sensitive to delays. Therefore the additional consideration of maximization of robustness should be taken into account. The resulting (large scale) bicriteria set partitioning problems can be solved by exact methods using state of the art integer programming techniques [4]. The planning of railway network infrastructure capacity has the goals of maximizing the number of trains that can use the infrastructure element (e.g. a station) and to maximize robustness of the solution to disruptions in operation. This problem can be modelled as (again large scale) set packing problem with two objectives [19]. Heuristic methods are currently used for its solution. Other recent applications include exact and heuristic methods for portfolio optimization, e.g. [7], a heuristic method for multiobjective vehicle routing problems [29], telecommunication networks [81] and timetabling problems [9].

2 Definitions

A multiobjective combinatorial optimization problem can be defined as follows. Given a finite set $A = \{a_1, \ldots, a_n\}$ a subset $X \subseteq 2^A$ defines a feasible set with a combinatorial structure. Objective functions are obtained from weight functions $w_j : A \to \mathbb{Z}$, $j = 1, \ldots, Q$ by defining for $S \in X$ $z^j(S) = \sum_{a \in S} w_j(a)$ (sum objective) or $z^j(S) = \max_{a \in S} w_j(a)$ (bottleneck objective). A multiobjective combinatorial optimization problem is then

$$\text{“}\min_{S \in X}\text{”}(z^1(S), \ldots, z^Q(S)). \qquad \text{(MOCO)}$$

The definition of “min” and thus the definition of an optimal solution of (MOCO) depends on the order of $\mathbb{R}^Q$. In *Pareto optimality (efficiency)* $S \in X$ is called Pareto optimal (efficient) if there is no $S' \in X$ with $z^j(S') \leq z^j(S)$, $j = 1, \ldots, Q$ and $z^q(S') < z^q(S)$ for some q. In this case $z(S) = (z^1(S), \ldots, z^Q(S))$ is called efficient (non-dominated) and the set of Pareto optimal (efficient) solutions is denoted by E. *Lexicographic optimality* is defined with respect to the lexicographic order $z(S_1) <_{lex} z(S_2)$ if $z^j(S_1) < z^j(S_2)$ and j is the smallest index such that $z^j(S_1) \neq z^j(S_2)$. It is possible to consider lexicographic optimality with respect to one or all permutations of the objective functions z^j. For *max-ordering optimality* the goal is to minimize the worst objective function, i.e. $\min_{S \in X} \max_{j=1,\ldots,Q} z^j(S)$. *Lexicographic max-ordering optimality* considers the vectors of objective values $z(S)$ reordered non-increasingly and compares these reordered vectors lexicographically. Because of the combinatorial structure a feasible solution $S \in X$ can be represented as a binary vector $x \in \{0,1\}^n$ by defining $x_i = 1$ if and only if $a_i \in S$, and 0 otherwise. Thus, (MOCO) is a discrete optimization problem, with n variables x_i, $i = 1, \ldots, n$, m constraints of specific structure defining X, Q objectives z^j, $j = 1, \ldots, Q$, and an order of $\mathbb{R}^Q$ to define optimality. In this paper we will be mainly concerned with the Pareto optimality concept.

3 Characteristics of MOCO Problems

3.1 Supported and Nonsupported Efficient Solutions

The most important property of (MOCO) can be explained via scalarization using convex combinations of objectives. A multiobjective linear programme (MOLP) is the problem $\min\{Cx : Ax = b, x \geq 0\}$, where C is a $Q \times n$ objective function matrix. A fundamental result in multiobjective linear programming is that E is the set of solutions of parametric linear programmes $\min\{\sum_{j=1,\ldots,Q} \lambda_j c^j x : Ax = b, x \geq 0\}$ with $0 < \lambda_j < 1$ and $\sum_{j=1}^{Q} \lambda_j = 1$. The non-convexity of the feasible set of a MOCO problem, however, implies that supported efficient solutions SE (solutions of parametric problems, as in

(MOLP)), as well as nonsupported efficient solutions NE exist. This is even true for problems in which the convex hull of feasible solutions coincides with the feasible set of the LP relaxation (implying that total unimodularity is not as useful for MOCO as it is in single objective combinatorial optimization). Adding to the difficulty is the number of efficient solutions. Theoretical results show that E might be exponential in problem size, in fact every feasible solution might be efficient. Such problems are clearly intractable in terms of polynomial time algorithms. Problems for which this behaviour has been shown include spanning tree [42], shortest path [45], travelling salesperson [30]. Even the set of supported solutions SE can be exponential in problem size (network flow problems [70]). Experimental solutions reveal a more differentiated picture. For knapsack problems the number of supported solutions grows linearly, the number of nonsupported solutions grows exponentially [87]. It also seems to be the case that the numerical values of the objectives have an impact on the number of efficient solutions and the size of SE/NE [18]. The situation is better for bottleneck objectives, see e.g. [62].

3.2 Computational Complexity

The existence of nonsupported efficient solutions already indicates that MOCO problems are hard. For a more thorough investigation we have to define a decision problem related to (MOCO): *Given* $k_1, \ldots, k_Q \in \mathbb{Z}$ *does there exist some* $S \in X$ *such that* $z^j(S) \leq k_j,\ j = 1, \ldots, Q$*?* Closely related is the counting problem: *How many* $S \in X$ *satisfy* $z^j(S) \leq k_j,\ j = 1, \ldots, Q$*?* Research results indicate that decision versions of MOCO problems are “always” $\mathbb{NP}$-complete and the counting versions often $\#\mathcal{P}$-complete. The following problems are among those known to be $\mathbb{NP}$-complete: the unconstrained (MOCO) [20], multiobjective shortest path [74], multiobjective spanning tree [10] and multiobjective assignment [74]. The proofs show that knapsack or partition structures are present in these problems. In addition, all single objective $\mathbb{NP}$-hard problems are obviously $\mathbb{NP}$-hard in the multiobjective case. We briefly summarize results for other optimality concepts. The max-ordering problem with sum objectives is $\mathbb{NP}$-hard in general [11]. The max-ordering problem with bottleneck objectives is as easy or difficult as the single objective counterpart [21]. Lexicographic problems are often easy (for a given permutation of the objectives), because the lexicographic order is a total order.

4 Exact Solution Methods

4.1 Weighted Sums Method

The most popular albeit not really appropriate method for solving (MOCO) problems and multiobjective programmes in general is the weighted sums method. The scalarized problem

$$\min \left\{ \sum_{j=1}^{Q} \lambda_j z^j(x) : x \in X \right\} \qquad (P_\lambda)$$

has to be solved for all $\lambda \in I\!R^Q$ with $0 \leq \lambda_j \leq 1$ and $\sum_{j=1}^{Q} \lambda_j = 1$. The method finds all supported efficient solutions, but of course no unsupported ones. In early papers on MOCO it is striking that nonsupported efficient solutions have not been considered, presumably because their existence was not known. The weighted sums method is most often used when $Q = 2$, a generalization for $Q \geq 3$ is not straightforward and no general technique is known. Applications include assignment [17], knapsack [69], shortest path [88], spanning tree [42], etc.

4.2 Compromise Programming

The idea of compromise programming is to minimize the distance to the ideal point z^I defined by $z_j^I := \min_{x \in X} z^j(x)$. Most often a Tchebycheff norm is used as distance measure, so that the compromise program becomes

$$\min \left\{ \max_{j=1}^{Q} \{\lambda_j |z^j(x) - z_j^I|\} : x \in X \right\}. \qquad \text{(CP)}$$

With appropriate choices of λ all efficient solutions can be found. The drawback, however, is that (CP) is usually $I\!N\!P$-hard (shortest path [64]). Note that if the Tchebycheff norm is replaced by the l_1 norm (CP) coincides with (P_λ). With the l_p norm, $1 < p < \infty$, (CP) has a nonlinear objective, a problem which is hardly ever considered, a rare exception is [85]. Also note that because problems of similar form as (CP) are often used in interactive methods, the $I\!N\!P$-hardness results cast some shadow on the effectiveness of interactive procedures in multiobjective combinatorial optimization.

4.3 ε-Constraint and Elastic Constraint Method

The main idea of these methods is to minimize only one of the objectives whilst imposing constraints on the others. The scalarization used in the ε-constraint method is

$$\min \left\{ z^i(x) : x \in X, z^j(x) \leq \varepsilon_j, j \neq i \right\}. \qquad (\varepsilon\text{C})$$

It is possible to find all Pareto optimal solutions, but in a (MOCO) context the problem (εC) where $z^i(x)$ are sum objectives is often $I\!N\!P$-hard because of the (knapsack) constraints on the objectives. In the literature it is mostly used for bottleneck objectives, e.g. for assignment, knapsack, spanning tree, TSP [62]. In this case it is a very effective method.

The elastic constraints method can be seen as a modification of the original ε-constraint method based on the idea to reduce the computational difficulties created by the constraints by making them elastic, i.e. the scalarization becomes

$$\min \left\{ z^i(x) + \sum_{j \neq i} p_j su_j : x \in X, z^j(x) - sl_j + su_j = \varepsilon_j, j \neq i \right\}, \qquad \text{(EC)}$$

where sl_j and su_j are slack and surplus variables for the constraints on the objectives. The method is also able to find all Pareto optimal solutions and in addition shows computationally superior performance in hard but structured combinatorial problems (set partitioning in [4]). Interestingly, the method is a common generalization of both the weighted sums and ε-constraint methods.

4.4 Ranking

In combinatorial optimization the ranking of solutions, or the computation of K-best solutions, has received considerable attention. This concept can be exploited for finding efficient solutions of (MOCO) problems. For problems with two objectives the Nadir point z^N is defined as $z_j^N := \min_{x \in X} \{z^j(x) : z^i(x) = z_i^I, i \neq j\}$. Then, because z^I, z^N are lower and upper bounds on efficient solutions the following procedure is possible: Start by finding a solution with $z^1(x) = z_1^I$ and continue to find second best, third best, ..., K-best solutions with respect to z^1 until the value z_1^N is reached. Algorithms based on this idea have been used to solve shortest path [60] problems. The idea of ranking is also useful for max-ordering even in the general case of $Q > 2$ [26,42]. To properly generalize the ranking approach to more than three objectives the consideration of level sets of the objectives is currently under investigation [28].

4.5 Specific Methods

Researchers have also pursued the path of generalizing specific methods for solving particular single objective combinatorial problems to the multicriteria case. These efforts resulted in work on multiobjective dynamic programming which is based on a recursion formula $\min \left(g_N(x_n) + \sum_{k=0}^{N-1} g_k(x_k, u_k) \right)$ with a vector cost function g, state variables x_k, and control variables u_k. Naturally, this research has focused on problems for which dynamic programming formulations have been successfully applied in the single objective case, such as shortest path problems, e.g. [54] and knapsack problems, e.g. [52]. Other specific methods include label correcting methods for shortest path problems [59] and greedy algorithms for spanning tree problems [1].

4.6 Two Phases Method

To conclude this section, we describe a method that is generic for the MOCO area. Its name illustrates the main idea: In Phase 1, find all supported efficient solutions and use this information in Phase 2 to generate nonsupported efficient solutions. This information can be reduced costs, bounds etc. The method performs particularly well if the single objective counterpart is polynomially solvable, so that solution of each (P_λ) problem is "easy". So far it has been applied to a number of biobjective problems: network flow [55], assignment [84], spanning tree [67], knapsack [87]. A generalization to more objectives is still an open question, due to the same reasons the weighted sums approach for $Q \geq 3$ is still not definitively settled.

5 Heuristic Solution Methods

5.1 Approximation in a Multiobjective Context

The challenge for heuristic methods in multiobjective programming is that rather than finding one "good" solution the objective value of which approximates the optimal solution value of the problem, we have to approximate the unknown set E. Multiple objective heuristics (MOH) methods have to provide a good tradeoff between the quality of the set of potential efficient solutions $\widehat{E}$ and the time and memory requirements. When the method refers to a metaheuristic one talks about multiple objective metaheuristic (MOMH). From a historical perspective, metaheuristic techniques for the solution of multiobjective problems have appeared since 1984, in the following order: Genetic Algorithms (1984) [73], Neural Networks (1990) [58], Simulated Annealing (1992) [75], and Tabu Search (1996) [35]. Even though it was easy to classify the pioneer methods as either evolutionary algorithms or neighborhood search algorithms, they are often hybridized today. A central question concerns the quality of a set of potential efficient solutions. Various researchers have contributed to the discussion of how to measure it. These contributions can be divided into those that consider the case when E is known [83] and include criteria of coverage, uniformity, and cardinality [71] or integrated convex preference [51]. The other broad group are those that consider comparison of approximations, such as evaluations of approximations [43] and metrics of performance [89] or the comparison with bounds and bound sets [23]. Considering the number of recent publications, approximation methods in multiobjective programming receive more and more attention. The following discussion is restricted to MOMH designed to identify sets of potential efficient solutions $\widehat{E}$ for MOCO problems.

5.2 Evolutionary Algorithms

Evolutionary methods manage a population of solutions rather than a single feasible one. In general, they start from an initial population and combine principles of self adaptation, i.e. independent evolution, and cooperation, i.e. the exchange of information between individuals, for improving solution quality. Thus, they develop a parallel process where the whole population contributes to the evolution process to generate $\widehat{E}$. The first multiobjective evolutionary algorithm (MOEA) was the Vector Evaluated Genetic Algorithm (VEGA) by Schaffer [72]. For each generation three stages are performed. The population is divided into Q subpopulations S^q according to performance in objective q. Subpopulations are then shuffled to create a mixed population. Genetic operators such as mutation and crossover are applied producing new potential efficient individuals. This process is repeated for N_{gen} iterations. The approximations achieved with VEGA typically showed good performance towards the extremes (close to optimality for individual objectives) but poor quality for areas of E corresponding to compromise solutions. Methods of ranking, niching and sharing have been proposed later to have a uniform convergence an distribution of individuals along the efficient frontier. The idea of ranking methods [40] is to subdivide the population into groups of different ranks according to their quality. Niches are neighbourhoods of solutions in objective space centered at candidate solutions and with some radius σ_{sh}. Based on the number N of solutions in these niches the selection of individuals can be influenced to areas in which niches are sparsely populated to aim at greater uniformity of distribution along the efficient frontier. A number of important implementations of MOEA have been published in recent years, there are even a number of surveys on the topic (see [12,13,33,50]). Here we describe the methods which have been used for (MOCO).

- Pioneer MOEAs: Vector Evaluated Genetic Algorithm by Schaffer, 1984 [72]; Multiple Objective Genetic Algorithm by Fonseca and Fleming, 1993 [32]; Nondominated Sorting Genetic Algorithm by Srinivas and Deb, 1994 [77]; Niched Pareto Genetic Algorithm by Horn, Nafpliotis and Goldberg, 1994 [47].
- Multiple Objective Genetic Algorithm (MOGA) by Murata and Ishibuchi, 1995 [63]. This method is based on a weighted sum of objective functions to combine them into a scalar fitness function using weight values generated randomly in each iteration. Later they coupled a local search with genetic algorithm, introducing the memetic algorithm principle for multiobjective problems.
- Morita's method (MGK) by Morita, Gandibleux and Katoh, 1998 [36]. Seeding solutions, i.e. greedy or supported solutions, are put in the initial population to initialise the algorithm with good genetic information. The biobjective knapsack problem is used to validate the principle. It becomes a memetic algorithm when a local search is performed on each new potential efficient solution [37].
- Strength Pareto Evolutionary Algorithm (SPEA) by Zitzler and Thiele, 1998 [90]. SPEA takes the best features of previous MOEAs and includes

them in a single algorithm. The multiobjective multi-constraint knapsack problem has been used as benchmark to evaluate the method [91].

- Multiple Objective Genetic Local Search (MO-GLS) by Jaszkiewicz, 2001 [49]. This method hybridizes recombination operators with local improvement heuristics. A scalarizing function is drawn at random for selecting solutions, which are recombined and their offspring are improved using heuristics.
- Multiple Objective Genetic Tabu Search (MOGTS) by Barichard and Hao, 2002 [4]. Another hybrid method where a genetic algorithm is coupled with a tabu search. MOGTS has been evaluated on the multi-constraint knapsack problem.

5.3 Simulated Annealing Based Metaheuristics

In 1992, Serafini has published the first ideas about multiobjective simulated annealing [75] in a multiobjective context. At the same time, Ulungu introduced MOSA [83], one of the most popular simulated annealing based methods. It is a direct derivation of the simulated annealing principle to deal with multiple objectives. Starting from an initial solution x_0 and a neighbourhood structure $\mathcal{N}(x_0)$, MOSA computes approximations using a weight set Λ defining search directions $\lambda \in \Lambda$ and a local aggregation mechanism $S(z(x), \lambda)$ together with a cooling schedule to accept deteriorations in values with decreasing probability. Like all neighbourhood search based methods, MOSA combines several sequential processes in the objective space Z. For each λ in a set of weights Λ it starts with a randomly generated solution x. Then a solution in the neighbourhood of x is generated and accepted if it is either better (dominates x) or based on a probability depending on the current "temperature". Next the set of potential efficient solutions $\widehat{E_\lambda}$ in direction λ and other parameters are updated. The search stops after a certain number of iterations or when a predetermined temperature is reached. Finally the sets $\widehat{E_\lambda}$ are merged. Multiobjective metaheuristics based on simulated annealing published in the literature are the following.

- Multiobjective Simulated Annealing (MOSA) by Ulungu, 1993 [83].
- Engrand's method, 1997 [31] revised by Park and Suppapitnarm [66]. The method uses only the non-domination definition to select potential efficient solutions, avoiding the management of search direction and aggregation mechanism.
- Pareto Simulated Annealing (PSA) by Czyzak and Jaszkiewicz, 1998 [15]. PSA also uses a weighted sum. However, a sample set of initial solutions $S \subset X$ is combined with an exploration principle exploiting interaction between solutions to guide the generation process through the values of λ.
- Nam and Park's method, 2000 [65]. Another simulated annealing based method. The authors show good results on comparison with MOEA.

- Other simulated Annealing based methods. Bicriteria scheduling problems on a single machine [53]; Interactive SA-TS hybrid method for 0-1 multiobjective problems [3]; Trip planning problem [39]; Aircrew rostering problem [57]; Assembly line balancing problem with parallel workstations [61]; Analogue filter tuning [82].

5.4 Tabu Search Based Metaheuristics

Extensions of tabu search to multiobjective programming are recent in comparison with other classical metaheuristics. The first methods use a tabu process guided automatically by the current approximation obtained [35] or by a decision-maker in an interactive way [78]. These methods start from an initial solution x_0, use a neighbourhood structure $\mathcal{N}(z(x_0))$ and search directions λ. The tabu process with its memory structure is applied with a local aggregation mechanism $s(z(x), z^U, \lambda)$ that involves a reference point z^U to browse the objective space. Hybrid methods appeared a short time later, trying to improve the diversification of solutions along the efficient frontier. Ideas come from MOEA, like the use of a population [44], or a combination of tabu search with genetic algorithms [1]. Multiobjective tabu search procedures have been applied mainly on MOCO problems, especially on the knapsack problem. In the literature one can find the following MOMH based on tabu search.

- "False MOMH" using tabu search. They are not designed to reach a (sub)set of potential efficient solutions. (MOCO) is solved through a sequence of Q single objective problems with penalty terms [46], or through solution of (P_λ) [16].
- Multiobjective Tabu Search (MOTS) by Gandibleux, Mezdaoui and Fréville, 1997 [35]. The method has been tested on an unconstrained permutation problem, and later on the biobjective knapsack problem [34] using bounds to reduce the search space.
- Sun's method, 1997 [78]. This is an interactive procedure using a tabu search process as solver of combinatorial optimization subproblems. The components used to design the tabu search process are almost the same than in MOTS [35]. The method has been used for facility location planning [2].
- Multiobjective Tabu Search (MOTS*) by Hansen, 1997 [44]. This method uses a generation set (i.e. a number of solutions rather than one, each of which has its own tabu list) and a drift criterion. Results are available for the knapsack problem, and also for the resource constrained project scheduling problem [86].
- Ben Abdelaziz, Chaouachi and S. Krichen's hybrid method, 1999 [1]. The authors present a mutiobjective hybrid heuristic for the knapsack problem. The method is a mix of tabu search and a genetic algorithm.

- Baykasoglu, Owen and Gindy's method, 1999 [6]. Another tabu search based method designed to handle any type of variable. The method has been also used for goal programming problems [5].
- Other tabu search based methods have been developed for scheduling problems [56] and the trip planning problem [39].

5.5 Other Methods

Besides these multiobjective versions of now classical metaheuristic methods there exist other MOMH. We are aware of Artificial Neural Networks ANN [58,79,80], Greedy Randomized Adaptive Search Procedure GRASP [38], Ant Colony Systems ACO [41,48,76], and Scatter Search [8].

6 Directions of Research and Resources

The state of the art in multiobjective combinatorial optimization indicates a number of directions of research that are promising and should be considered to make substantial progress in the field. We list some of these here, divided into theory, methods, and applications. In the theory of MOCO an interesting question is which results in single objective combinatorial optimization are still valid when $Q > 1$? E.g. the Martello and Toth bound for knapsack problems is not valid when $Q = 2$. Further investigation into bound sets (started in [23]) and Nadir points (see [27]) can be expected to lead to better methods. In terms of the hardness of MOCO problems the question of whether there are easy and hard problems in MOCO in a sense other than $\mathbb{NP}$-hardness arises. The quality of approximations and the representation of Pareto sets by smaller subsets are exciting topics for research. As far as methods are concerned we point out that exact methods for $Q \geq 3$ objectives are not available. A closer look at the two phases method for $Q = 2$ when the single objective problem is $\mathbb{NP}$-hard should provide better understanding of MOCO. In the area of heuristics a fundamental question is the performance of generic MOMH versus problem specific MOMH. Also, the effectiveness of MOMH for different problems should be considered, or the use of semi-exact methods that may use bounds to reduce search space as in [34] is promising. For applications there is the general question of the choice between methods that generate the efficient set as opposed to interactive methods. Can guidelines for this choice be developed? The study of real world problems as MOCO models is becoming increasingly important. In this context we note that practical MOCO problems should not be treated as single objective problems, as has often been the case in the past. For further references and a more detailed exposition of the topics of this paper we refer to the publications [22,24]. Also, a library of numerical instances of MOCO problems is available on the internet. At the time of printing the library includes instances for the multiobjective assignment, knapsack, set covering, set packing, and traveling salesman problems, as well as test Problems for multiobjective optimizers. The library is located at `www.terry.uga.edu/mcdm/`.

References

1. F. Ben Abdelaziz, J. Chaouachi, and S. Krichen. A hybrid heuristic for multiobjective knapsack problems. In S. Voss et al., editors, *Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization*, pages 205–212. Kluwer, Dordrecht, 1999.
2. P. Agrell, M. Sun, and A. Stam. A tabu search multi-criteria decision model for facility location planning. In *Proceedings of the 1997 DSI Annual Meeting*, 2:908–910. Atlanta, 1997.
3. M.J. Alves and J. Climaco. An interactive method for 0-1 multiobjective problems using simulated annealing and tabu search. *J. Heuristics*, 6(3):385–403, 2000.
4. V. Barichard and J.K. Hao. Un algorithme hybride pour le problème de sac à dos multi-objectifs. Huitièmes Journées Nationales sur la Résolution Pratique de Problèmes NP-Complets JNPC'2002, Nice, France, 27–29 May 2002.
5. A. Baykasoglu. MOAPPS 1.0: Aggregate production planning using the multiple objective tabu search. *Int. J. Prod. Res.*, 39(16):3685–3702, 2001.
6. A. Baykasoglu, S. Owen, and N. Gindy. A taboo search based approach to find the Pareto optimal set in multiple objective optimisation. *J. Eng. Optim.*, 31:731–748, 1999.
7. J.E. Beasley, T.J. Chang, N. Meade, and Y.M. Sharaiha. Heuristics for cardinality constrained portfolio optimisation. *Comput. Oper. Res.*, 27(13):1271-1302, 2000.
8. R. Beausoleil. Multiple criteria scatter search. In *MIC'2001 - 4th Metaheuristics International Conference*, pages 539–543. Porto, Portugal, July 16-20, 2001.
9. E.K. Burke, Y. Bykov, and S. Petrovic. A multi-criteria approach to examination timetabling. In E.K. Burke and W. Erben, editors, *The Practice and Theory of Automated Timetabling III, Lect. Notes Comput. Sci.* 2079:118–131. Springer, Berlin, 2000.
10. P.M. Camerini, G. Galbiati, and F. Maffioli. The complexity of multi-constrained spanning tree problems. In L. Lovasz, editor, *Theory of Algorithms*, pages 53 – 101. North-Holland, Amsterdam, 1984.
11. S. Chung, H.W. Hamacher, F. Maffioli, and K.G. Murty. Note on combinatorial optimization with max-linear objective functions. *Discrete Appl. Math.*, 42:139–145, 1993.
12. C.A. Coello. A comprehensive survey of evolutionary-based multiobjective optimization techniques. *Knowledge and Information Systems*, 1(3):269–308, 1999.
13. C.A. Coello. An updated survey of GA-based multiobjective optimization techniques. *ACM Computing Surveys*, 32(2):109–143, 2000.
14. H.W. Corley. Efficient spanning trees. *J. Optim. Theory Appl.*, 45(3):481–485, 1985.
15. P. Czyzak and A. Jaszkiewicz. Pareto simulated annealing – A metaheuristic technique for multiple objective combinatorial optimization. *J. Multi-Criteria Decis. Anal.*, 7(1):34–47, 1998.

16. G. Dahl, K. Jörnsten, and A. Lokketangen. A tabu search approach to the channel minimization problem. In *Proceedings of the International Conference on Optimization Techniques and Applications (ICOTA'95)*, 369–377. World Scientific, Singapore, 1995.
17. H.M. Dathe. Zur Lösung des Zuordnungsproblems bei zwei Zielgrößen. *Z. Oper. Res.*, 22:105–118, 1978.
18. F. Degoutin and X. Gandibleux. Un retour d'expérience sur la résolution de problèmes combinatoires bi-objectifs. Programmation Mathématique Multi-Objectif PM2O V Meeting, Angers, France, 17 May 2002.
19. X. Delorme, J. Rodriguez, and X. Gandibleux. Heuristics for railway infrastructure saturation. In *ATMOS 2001 Proceedings. Electronic Notes in Theoretical Computer Science* 50:41–55. URL: http://www.elsevier.nl/locate/entcs/volume50.html. Elsevier Science, Amsterdam, 2001.
20. M. Ehrgott. Approximation algorithms for combinatorial multicriteria optimization problems. *Int. Transac. Oper. Res.*, 7:5–31, 2000.
21. M. Ehrgott. *Multiple Criteria Optimization – Classification and Methodology.* Shaker, Aachen, 1997.
22. M. Ehrgott and X. Gandibleux. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR Spektrum*, 2000.
23. M. Ehrgott and X. Gandibleux. Bounds and bound sets for biobjective combinatorial optimization problems. In M. Köksalan and S. Zionts, editors, *Multiple Criteria Decision Making in the New Millenium, Lect. Notes Econ. Math. Syst.* 507:242–253. Springer, Berlin, 2001.
24. M. Ehrgott and X. Gandibleux, editors. *Multiple Criteria Optimization – State of the Art Annotated Bibliographic Surveys*, volume 52 of *Kluwer's International Series in Operations Research and Management Science.* Kluwer, Norwell, 2002.
25. M. Ehrgott and D.M. Ryan. Constructing robust crew schedules with bicriteria optimization. *J. Multi Criteria Decis. Anal.* in print, 2003.
26. M. Ehrgott and A.J.V. Skriver. Solving biobjective combinatorial max-ordering problems by ranking methods and a two-phases approach. *Eur. J. Oper. Res.*, in print, 2003.
27. M. Ehrgott and D. Tenfelde-Podehl. Computation of ideal and Nadir values and implications for their use in MCDM methods. *Eur. J. Oper. Res.*, in print, 2003.
28. M. Ehrgott and D. Tenfelde-Podehl. A level set method for multiobjective combinatorial optimization: Application to the quadratic assignement problem. Technical report, Universität Kaiserslautern, 2002.
29. N. El-Sherbeny. *Resolution of a vehicle routing problem with a multi-objective simulated annealing method.* PhD thesis, Université de Mons-Hainaut, 2001.
30. V.A. Emelichev and V.A. Perepelitsa. On cardinality of the set of alternatives in discrete many-criterion problems. *Discrete Mathematics and Applications*, 2(5):461–471, 1992.
31. P. Engrand. A multi-objective approach based on simulated annealing and its application to nuclear fuel management. In *Proceedings of the 5th ASME/SFEN/JSME International Conference on Nuclear Engineering. Icone 5, Nice, France 1997*, pages 416–423. American Society of Mechanical Engineers, New York, 1997.

32. C.M. Fonseca and P.J. Fleming. Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In S. Forrest, editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 416–423. Morgan Kaufman, San Francisco, 1993.
33. C.M. Fonseca and P.J. Fleming. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary Computation*, 3(1):1–16, 1995.
34. X. Gandibleux and A. Fréville. Tabu search based procedure for solving the 0/1 multiobjective knapsack problem: The two objective case. *J. Heuristics*, 6(3):361–383, 2000.
35. X. Gandibleux, N. Mezdaoui, and A. Fréville. A tabu search procedure to solve multiobjective combinatorial optimization problems. In R. Caballero, F. Ruiz, and R. Steuer, editors, *Advances in Multiple Objective and Goal Programming, Lect. Notes Econ. Math. Syst.* 455:291–300. Springer, Berlin, 1997.
36. X. Gandibleux, H. Morita, and N. Katoh. A genetic algorithm for 0-1 multiobjective knapsack problem. In *International Conference on Nonlinear Analysis and Convex Analysis (NACA98) Proceedings, July 28-31 1998, Niigata, Japan*, 4 pages, 1998.
37. X. Gandibleux, H. Morita, and N. Katoh. The supported solutions used as a genetic information in a population heuristic. In E. Zitzler et al., editors, *First International Conference on Evolutionary Multi-Criterion Optimization, Lect. Notes Comput. Sci.*, 1993:429–442. Springer, Berlin, 2001.
38. X. Gandibleux, D. Vancoppenolle, and D. Tuyttens. A first making use of GRASP for solving MOCO problems. Technical report, University of Valenciennes, France, 1998.
39. J.M. Godart. *Problèmes d'optimisation combinatoire à caractère économique dans le secteur du tourisme (organisation de voyages).* PhD thesis, Université de Mons-Hainaut, 2001.
40. D.E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, Reading, 1989.
41. M. Gravel, W.L. Price, and C. Gagné. Scheduling continuous casting of aluminium using a multiple-objective ant colony optimization metaheuristic. *Eur. J. Oper. Res.*, 143(1):218-229, 2002.
42. H.W. Hamacher and G. Ruhe. On spanning tree problems with multiple objectives. *Ann. Oper. Res.*, 52:209–230, 1994.
43. M. P. Hansen and A. Jaszkiewicz. Evaluating the quality of approximations to the non-dominated set. Technical report IMM-REP-1998-7, Technical University of Denmark, 1998.
44. M.P. Hansen. Tabu search for multiobjective combinatorial optimization: TAMOCO. *Control and Cybernetics*, 29(3):799–818, 2000.
45. P. Hansen. Bicriterion path problems. In G. Fandel and T. Gal, editors, *Multiple Criteria Decision Making Theory and Application, Lect. Notes Econ. Math. Syst.*, 177:109–127. Springer, Berlin, 1979.
46. A. Hertz, B. Jaumard, C. Ribeiro, and W. Formosinho Filho. A multi-criteria tabu search approach to cell formation problems in group technology with multiple objectives. *RAIRO - Rech. Opér.*, 28(3):303–328, 1994.
47. J. Horn, N. Nafpliotis, and D.E. Goldberg. A niched Pareto genetic algorithm for multiobjective optimization. In *Proceedings of the First IEEE Conference on Evolutionary Computation*, 1:82–87. IEEE Service Center, Piscataway, 1994.

48. S. Iredi, D. Merkle, and M. Middendorf. Bi-criterion optimization with multi colony ant algorithms. In E. Zitzler et al., editors, *First International Conference on Evolutionary Multi-Criterion Optimization*, *Lect. Notes Comput. Sci.*, 1993:359–372. Springer, Berlin, 2001.
49. A. Jaszkiewicz. Multiple objective genetic local search algorithm. In M. Köksalan and S. Zionts, editors, *Multiple Criteria Decision Making in the New Millennium*, *Lect. Notes Econ. Math. Syst.*, 507:231–240. Springer, Berlin, 2001.
50. D. Jones, S.K. Mirrazavi, and M. Tamiz. Multi-objective meta-heuristics: An overview of the current state-of-the-art. *Eur. J. Oper. Res.*, 137(1):1–9, 2002.
51. B. Kim, E.S. Gel, W.M. Carlyle, and J.W. Fowler. A new technique to compare algorithms for bi-criteria combinatorial optimization problems. In M. Köksalan and S. Zionts, editors, *Multiple Criteria Decision Making in the New Millenium*, *Lect. Notes Econ. Math. Syst.*, 507:113–123. Springer,Berlin, 2001.
52. K. Klamroth and M. Wiecek. A time-dependent single-machine scheduling knapsack problem. *Eur. J. Oper. Res.*, 135:17–26, 2001.
53. E. Koktener and M. Köksalan. A simulated annealing approach to bicriteria scheduling problems on a single machine. *J. Heuristics*, 6(3):311–327, 2000.
54. M.M. Kostreva and M.M. Wiecek. Time dependency in multiple objective dynamic programming. *J. Math. Anal. Appl.*, 173(1):289–307, 1993.
55. H. Lee and P.S. Pulat. Bicriteria network flow problems: Integer case. *Eur. J. Oper. Res.*, 66:148–157, 1993.
56. T. Loukil Moalla, J. Teghem, and P. Fortemps. Solving multiobjective scheduling problems with tabu search. In *Workshop on Production Planning and Control*, pages 18–26. Facultés Universitaires Catholiques de Mons, 2000.
57. P. Lučić and D.Teodorović. Simulated annealing for the multi-objective aircrew rostering problem. *Transportation Research A: Policy and Practice*, 33(1):19–45, 1999.
58. B. Malakooti, J. Wang, and E.C. Tandler. A sensor-based accelerated approach for multi-attribute machinability and tool life evaluation. *Int. J. Prod. Res.*, 28:2373, 1990.
59. E.Q.V. Martins. On a multicriteria shortest path problem. *Eur. J. Oper. Res.*, 16:236–245, 1984.
60. E.Q.V. Martins and J.C.N. Climaco. On the determination of the nondominated paths in a multiobjective network problem. *Methods Oper. Res.*, 40:255–258, 1981.
61. P.R. McMullen and G.V. Frazier. Using simulated annealing to solve a multiobjective assembly line balancing problem with parallel workstations. *Int. J. Prod. Res.*, 36(10):2717 - 2741, 1999.
62. I.I. Melamed and I.K. Sigal. A computational investigation of linear parametrization of criteria in multicriteria discrete programming. *Comp. Math. Math. Phys.*, 36(10):1341–1343, 1996.
63. T. Murata and H. Ishibuchi. MOGA: Multi-objective genetic algorithms. In *Proceedings of the 2nd IEEE International Conference on Evolutionary Computing*, pages 289–294. IEEE Service Center, Piscataway, 1995.
64. I. Murthy and S.S. Her. Solving min-max shortest-path problems on a network. *Nav. Res. Logist.*, 39:669–683, 1992.

65. D. Nam and C.H. Park. Multiobjective simulated annealing: A comparative study to evolutionary algorithms. *International Journal of Fuzzy Systems*, 2(2):87–97, 2000.
66. G. Parks and A. Suppapitnarm. Multiobjective optimization of PWR reload core designs using simulated annealing. In *Proceedings of the International Conference on Mathematics and Computation, Reactor Physics and Environmental Analysis in Nuclear Applications*, 2:1435–1444, Madrid, 1999.
67. R.M. Ramos, S. Alonso, J. Sicilia, and C. González. The problem of the optimal biobjective spanning tree. *Eur. J. Oper. Res.*, 111:617–628, 1998.
68. C. Reeves. *Modern Heuristic Techniques for Combinatorial Problems.* McGrawHill, London, 1995.
69. M.J. Rosenblatt and Z. Sinuany-Stern. Generating the discrete efficient frontier to the capital budgeting problem. *Oper. Res.*, 37(3):384–394, 1989.
70. G. Ruhe. Complexity results for multicriteria and parametric network flows using a pathological graph of Zadeh. *Z. Oper. Res.*, 32:59–27, 1988.
71. S. Sayin. Measuring the quality of discrete representations of efficient sets in multiple objective mathematical programming. *Math. Prog.*, 87:543–560, 2000.
72. J.D. Schaffer. *Multiple Objective Optimization with Vector Evaluated Genetic Algorithms.* PhD thesis, Vanderbilt University, Nashville, 1984.
73. J.D. Schaffer. Multiple objective optimization with vector evaluated genetic algorithms. In J.J. Grefenstette, editor, *Genetic Algorithms and their Applications: Proceedings of the First International Conference on Genetic Algorithms*, pages 93–100. Lawrence Erlbaum, Pittsburgh, 1985.
74. P. Serafini. Some considerations about computational complexity for multi objective combinatorial problems. In J. Jahn and W. Krabs, editors, *Recent advances and historical development of vector optimization, Lect. Notes Econ. Math. Syst.*, 294:222-232. Springer, Berlin, 1986.
75. P. Serafini. Simulated annealing for multiobjective optimization problems. In *Proceedings of the 10th International Conference on Multiple Criteria Decision Making, Taipei-Taiwan*, I:87–96, 1992.
76. P.S. Shelokar, S. Adhikari, R. Vakil, V.K. Jayaraman, and B.D. Kulkarni. Multiobjective ant algorithm for continuous function optimization: Combination of strength Pareto fitness assignment and thermo-dynamic clustering. *Found. Comp. Decis. Sci.*, 25(4):213–230, 2000.
77. N. Srinivas and K. Deb. Multiobjective optimization using non-dominated sorting in genetic algorithms. *Evolutionary Computation*, 2(3):221–248, 1994.
78. M. Sun. Applying tabu search to multiple objective combinatorial optimization problems. In *Proceedings of the 1997 DSI Annual Meeting*, 2:945–947. Atlanta, 1997.
79. M. Sun, A. Stam, and R. Steuer. Solving multiple objective programming problems using feed-forward artificial neural networks: The interactive FFANN procedure. *Manage. Sci.*, 42(6):835–849, 1996.
80. M. Sun, A. Stam, and R. Steuer. Interactive multiple objective programming using Tchebycheff programs and artificial neural networks. *Comput. Oper. Res.*, 27:601–620, 2000.
81. B. Thiongane, V. Gabrel, D. Vanderpooten, and S. Bibas. Le problème de la recherche de chemins efficaces dans un réseau de télécommunications. Francoro III, Québec, May 9-12, 2001.

82. M. Thompson. Application of multi objective evolutionary algorithms to analogue filter tuning. In E. Zitzler et al., editors, *First International Conference on Evolutionary Multi-Criterion Optimization*, *Lect. Notes Comput. Sci.*, 1993:546–559. Springer, Berlin, 2001.
83. E.L. Ulungu. *Optimisation combinatoire multicritère: Détermination de l'ensemble des solutions efficaces et méthodes interactives.* PhD thesis, Université de Mons-Hainaut, 1993.
84. E.L. Ulungu and J. Teghem. The two-phases method: An efficient procedure to solve bi-objective combinatorial optimization problems. *Found. Comput. Decis. Sci.*, 20(2):149–165, 1994.
85. A. Vainshtein. Vector shortest path problem in l_p norm. In *Simulation and Optimization of Complex Structure Systems*, pages 138–144. Omsk, 1987.
86. A. Viana and J. Pinho de Sousa. Using metaheuristics in multiobjective ressource constrained project scheduling. *Eur. J. Oper. Res.*, 120(2):359–374, 2000.
87. M. Visée, J. Teghem, M. Pirlot, and E.L. Ulungu. Two-phases method and branch and bound procedures to solve the bi-obective knapsack problem. *J. Glob. Optim.*, 12:139–155, 1998.
88. D.J. White. The set of efficient solutions for multiple objective shortest path problems. *Comp. Oper. Res.*, 9(2):101–107, 1987.
89. E. Zitzler, K. Deb, and L. Thiele. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary Computation*, 8(2):173–195, 2000.
90. E. Zitzler and L. Thiele. An evolutionary algorithm for multiobjective optimization: The strength Pareto approach. Technical report 43, Computer Engineering and Communication Networks Lab (TIK), Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, 1998.
91. E. Zitzler and L. Thiele. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, 1999.

Analysis of Trends in Distance Metric Optimisation Modelling for Operational Research and Soft Computing

D.F. Jones, M.Tamiz

Department of Mathematics, University of Portsmouth UK

Dylan.Jones@port.ac.uk, Mehrdad.Tamiz@port.ac.uk

Abstract

This paper provides a commentary and some analysis on recent advances in the field of distance metric optimisation, with particular reference to the place of distance metric optimisation within the overall disciplines of operational research and soft computing. The trend of integration and combination with other techniques is examined, with particular reference to the analytical hierarchy method, meta-heuristic methods, and data mining. Finally, some further thoughts on good modelling practice for distance metric optimisation models are offered.

Keywords : Distance metric optimisation, goal programming, analytical hierarchy process, meta-heuristic methods, data mining.

1. Introduction

Distance metric optimisation is characterized by the minimization of some distance function between the achieved levels of a set of objectives and either an ideal level or a decision maker desired level measured in terms of the same set of objectives. The well-known multi-objective techniques that fall into the category of distance metric optimization include goal programming, compromise programming, the reference-point method, and some interactive extensions of the previous methods. Mathematically speaking, the non-lexicographic distance metric optimisation minimisation function can be defined as:

$$Min\ z = \left[\sum_{i=1}^{q} \left[\frac{u_i n_i + v_i p_i}{k_i} \right]^{\rho} \right]^{1/\rho}$$

with an associated set of goals or objectives:

$$f_i(x) + n_i - p_i = b_i \quad i = 1,\ldots,q$$

and optional set of hard constraints:

$$x \in F$$

where x is the set of decision variables, $f_i(x)$ is a mathematical expression defining the achieved value of the i'th goal or objective, u_i and v_i are the weights associated with the penalization of the deviations (n_i, p_i respectively) from the desired or ideal level (b_i) of the i'th objective. A weight of zero associated with a deviation indicates the minimization of that deviation is unimportant to the decision maker. The term ρ is the distance metric used to measure the distance between the achieved and the desired or ideal levels of the set of objectives. Varying ρ between its end-point values of 1 and ∞ produces a range of solutions that vary between a ruthless optimization approach ($\rho = 1$) and a balanced approach that produces as equilibrated a solution as possible ($\rho = \infty$). The term k_i is a normalisation constant included to overcome incommensurability and hence to allow the deviations from the objectives to be compared directly. The traditional choice for the normalisation constant in compromise programming is the distance between the ideal and the nadir value for that objective, thus scaling all objectives onto a zero-one range. The anti-ideal value of the objective is sometimes used as a surrogate for the nadir value if the latter is too computationally difficult to compute. Popular normalisation methods for the goal programming model include the percentage, zero-one, and Euclidean methods. These are analysed by Tamiz and Jones [23] who also present an algorithm for measuring the level incommensurability and hence suggesting or automatically applying an appropriate normalisation technique.

This model covers all non-lexicographic distance metric optimisation techniques. This is sufficient to model compromise programming and non-pre-emptive (weighted) goal programming models. In order to extend the theory to other methods a lexicographic order must be introduced. This leads to the following algebraic formulation of the achievement function:

$$Min\,\mathrm{a} = \left[\sum_{i=1}^{q}\left[\frac{u_i^{(1)} n_i + v_i^{(1)} p_i}{k_i}\right]^{\rho_1}\right]^{1/\rho_1}, \left[\sum_{i=1}^{q}\left[\frac{u_i^{(2)} n_i + v_i^{(2)} p_i}{k_i}\right]^{\rho_2}\right]^{1/\rho_2},$$

$$\ldots\ldots, \left[\sum_{i=1}^{q}\left[\frac{u_i^{(L)} n_i + v_i^{(L)} p_i}{k_i}\right]^{\rho_L}\right]^{1/\rho_L}$$

where the commas represent the distinction between the L pre-emptive priority levels in the model. The distance metric used in the l'th priority level is given by ρ_l and the weights associated with penalisation of the negative and positive deviational variables of the i'th objective in the l'th priority level are given by $u_i^{(l)}$ and $v_i^{(l)}$ respectively. With the possibility of negative and zero weights, this model allows the lexicographic based distance optimisation models such as lexicographic goal programming and the reference point method to be modelled. Variations or partial variations of this model to allow various linear programming and distance-metric models to be formulated under a common framework are given by Romero [17], Ignizio [9], and Uria et el [24]. Romero, Tamiz, and Jones [18] propose further theoretical connections between the major techniques of distance metric optimisation. This topic is further developed by Ogryczak [16] and Ganjavi et al. [6].

The fundamentals and algebraic formulation of distance metric optimisation models have been outlined above. The remainder of this paper concentrates on the integration and combination of distance metric models with some other techniques within the Operational Research and Soft Computing disciplines. Section 2 details the interface of meta-heuristic methods and distance metric optimisation, section 3 of distance metric optimisation and the analytical hierarchy process, section 4 details the role of distance optimisation models in pattern classification, and section 5 offers some further thoughts and suggestions about good modelling practice in goal programming. The final section draws conclusions.

2. Distance Metric Optimisation and Meta Heuristic Methods

A meta-heuristic method draws on ideas and methodology from disciplines outside of artificial system optimization to provide algorithms for the solution of artificial system optimization models. Well-known meta heuristics include genetic algorithms, simulated annealing, and tabu search which draw on ideas from genetics, physics, and the social concept of Taboo respectively. Meta-Heuristic methods can be classified within the field of soft computing. The interface between meta-heuristic methods and the wider field of multi-objective programming, and in particular the use of genetic algorithm techniques for efficient frontier calculation, has been considerable. This can be traced to the fact that both genetic algorithms and Pareto frontier generation require a population of spaced solutions in order to work efficiently. A recent survey by Jones, Mirrazavi, and Tamiz [13] found that 90% of the journal articles related to multi-objective meta-heuristics are based around techniques for the calculation of the efficient set. The next most popular technique was goal programming, accounting for 7% of the articles, with compromise programming and interactive methods making up the remaining 3%. These statistics show that either the interface between distance metric optimization of meta-heuristics is non-existent in the sense of being of little benefit or is of practical benefit but has yet to be realized or developed. The discussion in the following paragraphs will argue in favour of the latter state of affairs.

In analyzing future developments in the interface between distance metric optimization and meta-heuristic methods three possible directions are apparent at this point in time. Firstly distance optimization techniques could be used to enhance the internal workings of the meta-heuristic method. This seems a possibility as there are various internal mechanisms in meta-heuristic techniques that rely on concepts of distance and deviation. The use of penalty functions [14] and of niching [8] in genetic algorithms fall into this category.

The second possible direction is to use the benefits of the meta-heuristic methods to provide enhanced or computationally faster solutions to certain distance metric optimization techniques. For example, genetic algorithm techniques have the potential to produce estimations of the compromise set in compromise programming in an analogous way to the methods that produce estimations of the efficient set in multi-objective programming. The commonality between the two methods would be the exploitation of the population-based nature of the genetic algorithm.

The third possible direction is the use of meta-heuristic methods to solve models that are too computationally complex or loosely defined to be modeled and solved using conventional means. This approach has proved very successful in the areas of single objective optimization and combinational optimization and the concepts can be transferred or modified to the distance optimization techniques. This is the most developed direction of the interface between meta-heuristic methods and distance metric optimization, particularly in respect to goal programming models. A recent goal programming survey [12] lists both simulated annealing and genetic algorithms as a solution tool for non-linear models in the field of engineering, and algorithms combining goal programming and Taboo search methods are available in the literature [1]. Mirrazavi, Jones, and Tamiz [15] present a decision support system capable of solving a wide variety of distance metric models by genetic algorithm means.

3. Distance Metric Optimisation and the Analytical Hierarchy Process

The analytical hierarchy process (AHP), developed by Saaty [19], has been one of the most widely used techniques in the field of decision analysis. The AHP framework allows for the determination of a set of priority weights from a matrix of pair-wise comparisons over the set of objectives given by the decision maker. These comparisons are made on a nine-point scale ranging from equal importance (1) to absolute importance (9).

The interface between distance metric optimisation and the analytical hierarchy process has been developed in two major directions. The first direction involves the use of a distance metric model as a surrogate to the standard Eigenvalue method in Saaty's original formulation. The earliest models of this type used the L_2 distance metric and were known as the least squares (LSM) and logarithmic least squares (LLSM) models, depending on whether the minimisation uses the logarithm of the matrix entries or not. The LLSM equates to the calculation of the geometric mean and hence demonstrates some good theoretical properties. Models based around the Logarithmic L_1 metric [2] and the L_∞ metric [5] have also been proposed. Islam, Biswal, and Alam [10] give an L_1 based method that incorporates interval judgements. Distance metric theory suggests that these solutions all form points in a compromise set corresponding to the metrics L_1, L_2, and L_∞ [25]. There is no reason why the intermediate distance-metric solutions corresponding to values of ρ other than 1,2, and ∞ should not also be considered.

The second direction in which the interface between distance metric optimisation and the AHP has been developed is that of the use of the AHP to set weights in a non preemptive goal programming model. This concept benefits both approaches as it quantifies the subjective weight setting of goal programming and provides an additional stage of analysis in a mathematical programming framework for the AHP. Gass [7] describes an early use of this approach in the context of large scale military planning. Jones and Tamiz [12] detail a range of reported applications of this combined method in computing and information technology, energy planning and production, environmental and waste management, health planning, management and strategic planning, and production planning.

4. Distance Metric Optimisation and Data Mining

Another area that has been a field of application of distance metric models is that of data mining. This field roughly involves the extraction and analysis of information from sets of data and falls into the general area of soft computing. The use of distance metric optimization in this field has concentrated around the use of constrained regression, regression with underlying distance metrics other than the standard $\rho = 2$ metric, and pattern classification and discriminant models. It is worth remembering that the original goal programming model [3] was introduced in the context of constrained regression in the context of an executive compensation model. Since then the theory of constrained regression has been developed, with the beneficial properties of least absolute value (LAV) regression being detailed by Sueyoshi and Sekitani[21]. This type of regression utilizes the $\rho = 1$ metric and therefore shows less sensitivity towards outliers than the other distance-metric models. Cooper, Lelas, and Sueyoshi [4] present an application of LAV regression to finance and further develop the theory to include the use of dual variables in the underlying goal programming model. The generalized regression model, using metrics from the compromise set from $\rho = 1$ through to $\rho = \infty$ is a possible future research direction, as it offers possibilities of a range of models with a parameter allowing the analyst to increase or decrease outlier sensitivity as necessary.

The area of pattern classification differs from regression analysis in that the task involved is to classify a set of observations into a number of well-defined groups based on their characteristics. The underlying problem here again involves a minimization of a form of distance function pertaining to either the number or amount of misclassifications across the set of observations [20]. The 'distance' of misclassification for an observation is that from the discriminant line that divides the classes in decision space and the multi-objective aggregation of the misclassified distances can be carried out using any metric from the compromise set between $\rho = 1$ and $\rho = \infty$. The case of using the number of misclassified observations as the measure of performance requires the use of mixed-integer programming techniques and is sometimes referred to as the $\rho = 0$ metric. A review of these so-called 'L_p norm methods' for pattern classification is given by Stam [20].

5. Some Further Observations on Goal Programming Modelling Practice

The traditional goal programming model is defined as having an achievement function comprised entirely of deviational variables[9]. This may take the form of a single weighted sum in weighted goal programming) or of a number of priority levels in lexicographic goal programming. In compromise programming models the normalised difference of the difference between the ideal and achieved values is minimized, this difference can be expressed as a deviational variable[18]. With the growing range of distance-metric model applications and integrations, there is more possibility of a mixed achievement function occurring. This case is defined by Jones [11] as a combination of decision and deviational variables in the achievement function. Assuming that the deviational and decision variables terms are separable then the weighted goal programming model achievement function can be written as:

$$Min\ z = \sum_{i=1}^{q} \frac{u_i n_i + v_i p_i}{k_i} + f(x).$$

This may cause problems with various types of solution and analysis such as Pareto efficiency detection and restoration [22] and also cause incommensurability in the model. In this case the following transformation is recommended:

$$Min\ z = \sum_{i=1}^{q+1} \frac{u_i n_i + v_i p_i}{k_i}$$

with the added constraint:

$$f(x) + n_{q+1} - p_{q+1} = 0$$

where $u_{q+1} = 0$, v_{q+1} is set to represent the relative importance of the minimisation of the term $f(x)$ to the decision maker, and k_{q+1} is set in order to give appropriate scaling to the term $f(x)$. The actual value of k_{q+1} is dependent on the type of normalisation used for the other objectives in the model, but one possibility is the Euclidean Norm of the coefficients of $f(x)$ [23].

This model can be considered better than the original formulation as it is more elegant and correct from a theoretical point of view; gives no hindrance to the use of solution and analysis techniques, and allows for correct scaling of all objectives in the model. It is also more convenient for integration with the other Operational Research and Soft Computing techniques described in this paper to have the goal programme expressed in standard form.

Hence it is recommended that any models with mixed achievement functions are transformed in this way.

6. Conclusions

This paper has analyzed some of the recent trends and applications in distance metric optimization. It has been shown that the subject, now approaching its fiftieth birthday since the conception of the goal programming model [3], continues to be relevant and applicable to the modern techniques being developed within the fields of operational research and soft computing. A substantial amount of change and development in order to apply the techniques of distance metric optimization to new and emerging application areas has taken place and needs to continue to take place. The distance metric framework laid down in the context of compromise programming by Yu [25] continues to offer a spectrum of possible solutions ranging between the pure optimization and balanced approaches, characterized by the $\rho = 1$ and $\rho = \infty$ metrics respectively, to each of these areas. This paper has been somewhat speculative in suggesting possible new research directions for each of these areas. The suggestions are not intended to be either definitive or exhaustive in terms of advances in distance metric optimization and its interface with the techniques detailed in the paper. They are, however, intended to demonstrate the fact that the area remains relevant and one in which there exist many avenues and interesting areas yet to be developed that should be of interest to both established and younger researchers with the field of multiple criteria decision making.

References

[1] Baykasoglu A (2001) Goal programming using multiple objective tabu search, *Journal Of The Operational Research Society*, **52** (12), 1359-1369

[2] Bryson. N (1995) A goal programming method for generating priority vectors, *Journal of the Operational Research Society*, **46** (5), 641-648.

[3] Charnes A, Cooper WW, Ferguson R (1955) Optimal estimation of executive compensation by linear programming, *Management Science*, **1**, 138-151.

[4] Cooper WW, Lelas V, Sueyoshi, T (1997) Goal programming models and their duality relations for use in evaluating security portfolio and regression relations, *European Journal of Operational Research,* **98**, 431-443.

[5] Despotis DK (1996) Fractional minmax goal programming: A unified approach to priority estimation and preference analysis in MCDM, Journal Of The Operational Research Society, **47** (8): 989-999.

[6] Ganjavi O, Aouni B, Wang Z (2002) Technical note on balanced solutions in goal programming, compromise programming and reference point method, *Journal Of The Operational Research Society*, **53** (8): 927-929

[7] Gass SI (1987) A process for determining priorities and weights for large scale goal programmes, *Journal of the Operational Research Society*, **37**, 779-785.

[8] Goldberg DE (1989) Genetic algorithms in search, optimization, and machine learning, *Addison Wesley Longman*, Reading, MA. Pages 185-197.

[9] Ignizio JP (1982) Linear programming in single and multiple objective systems, *Prentice Hall*, Englewood Cliffs, NJ.

[10] Islam R, Biswal MP, Alam SS (1997) Preference programming and inconsistent interval judgments, European Journal Of Operational Research, **97** (1): 53-62.

[11] Jones DF(1995) The design and development of an intelligent goal programming system, *Ph.D.Thesis,* University of Portsmouth, UK.

[12] Jones DF, Tamiz M (2002) Goal programming in the period 1990-2000, in *Multiple criteria*
optimization state of the art annotated Bibliographic surveys, M. Erghott X. Gandibleux (eds.) Kluwver.

[13] Jones DF., SK. Mirrazavi, and Tamiz M(2002) Multi-Objective meta-heuristics: an overview
of the current state-of-the art, *European Journal of Operational Research*, **137**, 1-9.[14] Michalewicz Z (1996) Genetic Algorithms + Data Structures = Evolution Structures, *Springer-Verlag*, Berlin, pages 80-93.

[15] SK Mirrazavi, DF Jones, and M. Tamiz (2003) MultiGen: an integrated multiple-objective solution system, to appear in *Decision Support Systems.*

[16] Ogryczak W (2001) Comments on Romero C, Tamiz M and Jones DF (1998). Goal programming, compromise programming and reference point method formulations: Linkages and utility interpretations, *Journal Of The Operational Research Society*, **52** (8): 960-962.

[17] Romero C (2001) Extended lexicographic goal programming: a unifying approach, Omega-International Journal Of Management Science, **29** (1): 63-71.

[18] Romero C, Tamiz M, Jones DF (1998) Goal programming, compromise programming and reference point method formulations: linkages and utility interpretations, Journal Of The Operational Research Society, **49** (9): 986-991.

[19] T.L. Saaty (1977) A Scaling Method for Priorities in Hierarchical Structures, *Journal of Mathematical Psychology*, 15 (3), 234-281.

[20] Stam A (1997) Nontraditional approaches to statistical classification: Some perspectives on L_p norm methods', *Annals of Operations Research,* **74**, 1-36.

[21] Sueyoshi, T, Sekitani K (1998) Mathematical properties of least absolute value estimation with serial correlation, *Asia-Pacific Journal of Operational Research*, **15**, 75-92.

[22] Tamiz, M, Jones, DF(1996) Goal Programming and Pareto Efficiency, *Journal of Information and Optimization Sciences,* **17**, 291-307.

[23] Tamiz, M and Jones DF (1997) An example of good modelling practice in goal programming: Means for overcoming incommensurability, Lecture notes in Economics and Mathematical Systems, R. Caballero, F. Ruiz(Eds), *Springer*, **455**, 29-37.

[24] Uria MVR, Caballero R, Ruiz F, Romero C (2002) Meta-goal programming, European Journal Of Operational Research, **136** (2): 422-429.

[25] Yu PL (1973) A class of solutions for group decision problems, *Management Science*, **19**, 936-946.

MOP/GP Approaches to Data Mining

Hirotaka Nakayama

Department of Information Science and Systems Engineering
Konan University
8-9-1 Okamoto, Higashinada, Kobe 658-8501, JAPAN
e-mail: nakayama@konan-u.ac.jp

Abstract. Recently, data mining is attracting researchers' interest as a tool for getting knowledge from data bases on a large scale. Although there have been several approaches to data mining, we focus on mathematical programming (in particular, multi-objective and goal programming; MOP/GP) approaches in this paper. Among them, Support Vector Machine (SVM) is gaining popularity as a method for machine learning. In pattern classification problems with two class sets, its idea is to find a maximal margin separating hyperplane which gives the greatest separation between the classes in a high dimensional feature space. This task is performed by solving a quadratic programming problem in a traditional formulation, and can be reduced to solving a linear programming in another formulation. However, the idea of maximal margin separation is not quite new: in 1960's the multi-surface method (MSM) was suggested by Mangasarian. In 1980's, linear classifiers using goal programming were developed extensively.

This paper presents a survey on how effectively MOP/GP techniques can be applied to machine learning such as SVM, and discusses their problems.

1 Introduction

One of main purposes in data mining is to discover knowledge in data bases with very large scale. Usually, machine learning techniques are utilized for this knowledge acquisition. Typical approaches to machine learning are 1) to find an explicit rule as *if-then rule* and 2) to judge newly observed data by an implicit rule which is usually represented as a nonlinear function. Well known ID3 (recently C5.0) and CART belong to the former category. On the other hand, artificial neural networks and mathemtical programming approaches belong to the latter category. In this paper, we focus on the latter category.

For convenience, we consider pattern classification problems. Let X be a space of conditional attributtes. For binary classification problems, the value of $+1$ or -1 is assigend to each data $\boldsymbol{x}_i$ according to its calss $\mathcal{A}$ or $\mathcal{B}$. The aim of machine learning is to predict which class newly observed data belong to on the basis of the given data set $(\boldsymbol{x}_i, y_i)$ $(i = 1, \dots, l)$, where $y_i = +1$ or -1.

For such a pattern classification problem, articial neural networks have been widely applied. However, the back propagation method is reduced to nonlinear optimization with multiple local optima, and hence difficult to apply to large scale problems. Another drawback in the back propagation method is in the fact that it is difficult to change the structure adaptively according to the change of environment in incremental learning. Rcently, Support Vector Machine (SVM, in short) is attracting interest of researchers, in particular, people who are engaged in mathematcal programming, because it is reduced to quadratic programming (QP) or linear programming (LP). One of main features in SVM is that it is a linear classifer with maximal margin on the feature space. The idea of maximal margin in linear classifier has a long history in mathematical programming and goal programming. In the following in this paper, we review it in brief and try to explain how effectively techniques in multi-objective programming and goal programming (MOP/GP) can be applied.

2 Multisurface Method (MSM)

Suppose that given data in a set X of n-dimensional Euclidean space belong to one of two categories $\mathcal{A}$ and $\mathcal{B}$. Let A be a matrix whose row vectors denote points in the category $\mathcal{A}$. Similarly, let B be a matrix whose row vectors denote points in the category $\mathcal{B}$. For simplicity of notation, we denote the set of points of $\mathcal{A}$ by A. The set of points of $\mathcal{B}$ is denoted by B similarly. MSM suggested by Mangasarian (1968) finds a piecewise linear discrimination surface separating two sets A and B by solving linear programming problems iteratively. The main idea is to find two hyperplanes parallel with each other which classify as many given data as possible:

$$g(\boldsymbol{w}) = \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{w} = \alpha$$
$$g(\boldsymbol{w}) = \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{w} = \beta$$

This is performed by the following algorithm:

Step 1 . Solve the following linear programming problem at k-th iteration (set $k = 1$ at the beginning):

$$\begin{array}{lll} \text{(MSM)} & \text{Maximize} & \phi_i(A,B) = \alpha - \beta \\ & \text{subject to} & \\ & & A\boldsymbol{w} \geqq \alpha\mathbf{1} \\ & & B\boldsymbol{w} \leq \beta\mathbf{1} \\ & & -\mathbf{1} \leqq \boldsymbol{w} \leqq \mathbf{1} \\ & & \boldsymbol{p}_i^T\boldsymbol{w} \geqq \frac{1}{2}\left(\frac{1}{2} + \boldsymbol{p}_i^T\boldsymbol{p}_i\right) \end{array} \tag{1}$$

where $\boldsymbol{p}_i$ is given by one of $\boldsymbol{p}_1^T = (\frac{1}{\sqrt{2}}, 0, \dots, 0)$, $\boldsymbol{p}_2^T = (-\frac{1}{\sqrt{2}}, 0, \dots, 0), \cdots \boldsymbol{p}_{2n}^T = (0, \dots, 0, -\frac{1}{\sqrt{2}})$.

Here the constraint (1) is introduced in order to avoid a trivial solution $\boldsymbol{w} = 0$, $\alpha = 0$, $\beta = 0$ from a linear approximation of $\boldsymbol{w}^T\boldsymbol{w} \geq \frac{1}{2}$. Namely,

$$\boldsymbol{w}^T\boldsymbol{w} \cong \boldsymbol{p}^T\boldsymbol{p} + 2\boldsymbol{p}^T(\boldsymbol{w} - \boldsymbol{p}) \geq \frac{1}{2}.$$

After soving LP problem (MSM) for each i such that $1 \leq i \leq 2n$, we take a hyperplane which classify correctly as many given data as possible. Let the solution be $\boldsymbol{w}^*, \alpha^*, \beta^*$, and let the corresponding value of objective function be $\phi^*(A, B)$.

If $\phi^*(A, B) > 0$, then we have a complete separating hyperplane $g(\boldsymbol{w}^*) = (\alpha^* + \beta^*)/2$. Set $\tilde{A}^k = \{\boldsymbol{x} \in X \mid g(\boldsymbol{w}^*) \geq (\alpha^* + \beta^*)/2\}$ and $\tilde{B}^k = \{\boldsymbol{x} \in X \mid g(\boldsymbol{w}^*) < (\alpha^* + \beta^*)/2\}$. $\tilde{A}^k$ and $\tilde{B}^k$ include the sets A and B in X, respectively, which is decided at this stage. Go to Step 3.

Otherwise, go to Step 2.

Step 2 . First, remove the points such that $\boldsymbol{x}^T\boldsymbol{w}^* > \beta^*$ from the set A. Let A^k denote the set of removed points. Take the separating hyperplne as $g(\boldsymbol{w}^*) = (\beta^* + \tilde{\beta})/2$ where $\tilde{\beta} = \text{Min}\ \{\boldsymbol{x}^T\boldsymbol{w}^* \mid \boldsymbol{x} \in A^k\}$. Let $\tilde{A}^k = \{\boldsymbol{x} \in X \mid g(\boldsymbol{w}^*) > (\beta^* + \tilde{\beta})/2\}$. The set $\tilde{A}^k$ denotes a subregion in the category $\mathcal{A}$ in X which is decided at this stage. Rewrite $X \backslash \tilde{A}^k$ by X and $A \backslash A^k$ by A.

Next, remove the points such that $\boldsymbol{x}^T\boldsymbol{w}^* < \alpha^*$ from the set B. Let B^k denote the set of removed points. Take the separating hyperplne as $g(\boldsymbol{w}^*) = (\alpha^* + \tilde{\alpha})/2$ where $\tilde{\alpha} = \text{Min}\ \{\boldsymbol{x}^T\boldsymbol{w}^* \mid \boldsymbol{x} \in B^k\}$. Let $\tilde{B}^k = \{\boldsymbol{x} \in X \mid g(\boldsymbol{w}^*) < (\alpha^* + \tilde{\alpha})/2\}$. The set $\tilde{B}^k$ denotes a subregion in the category $\mathcal{B}$ in X which is decided at this stage. Rewrite $X \backslash \tilde{B}^k$ by X and $B \backslash B^k$ by B.

Set $k = k + 1$ and go to Step 1.

Step 3. Construct a piecewise linear separating hypersurface for A and B by adopting the relevant parts of the hyperplanes obtained above.

Remark At the final p-th stage, we have the region of $\mathcal{A}$ in X as $\tilde{A}^1 \cup \tilde{A}^2 \cup \ldots \cup \tilde{A}^p$ and that of $\mathcal{B}$ in X as $\tilde{B}^1 \cup \tilde{B}^2 \cup \ldots \cup \tilde{B}^p$. Given a new point, its classification is easily made. Namely, since the new point is either one of these subregions in X, we can classify it by checking which subregion it belongs to in the order of $1, 2, \ldots, p$.

As stated above, if $\phi^*(A, B) > 0$, then the given data set can be linearly separated. Then, note that the parallel hyperplanes $\boldsymbol{g}(\boldsymbol{w}^*) = \alpha^*$ and $\boldsymbol{g}(\boldsymbol{w}^*) = \beta^*$ solving LP problem (MSM) provides a maximal margin.

3 Goal Programming Approaches to Pattern Classification

MSM often provides too complex discrimination boundaries, which results in a poor ability of generalization. In 1981, Freed-Glover suggested to get just a hyperplane separating two classes with as few misclassified data as possible by using goal programming (Freed-Glover (1981)). Let ξ_i denote the exterior deviation which is a deviation from the hyperplane of a point $\boldsymbol{x}_i$ improperly classified. Similary, let η_i denote the interior deviation which is a deviation from the hyperplane of a point $\boldsymbol{x}_i$ properly classified. Some of main objectives in this approach are as follows:

i) Minimize the maximum exterior deviation (decrease errors as much as possible)
ii) Maximize the minimum interior deviation (i.e., maximize the margin)
iii) Maximize the weighted sum of interior deviation
iv) Minimize the weighted sum of exterior deviation

Although many models have been suggested, the one considering iii) and iv) above may be given by the following linear goal programming:

$$\begin{array}{lll} \text{Minimize} & \sum_i^n (h_i\xi_i - k_i\eta_i) & \\ \text{subject to} & \boldsymbol{x}_i^T\boldsymbol{w} + b = \eta_i - \xi_i, & i \in I_A \\ & \boldsymbol{x}_i^T\boldsymbol{w} + b = -\eta_i + \xi_i, & i \in I_B \\ & \xi_i,\ \eta_i \geq 0 & i \in I_A \cup I_B \end{array}$$

Here, h_i and k_i are positive constants. It should be noted that the above formulation may yield some unacceptable solutions such as $\boldsymbol{w} = 0$ and unbounded solution. In order to avoid these unacceptable solutions, several normalization conditions have been suggested. For example, for some s

$$\boldsymbol{w}^T\mathbf{1} + b = s.$$

If the classification problem is linearly separable, then using the normalization $||\boldsymbol{w}|| = 1$, the separating hyperplane $H = \{\boldsymbol{x} \in R^m |\ \boldsymbol{w}^T\boldsymbol{x} + b = 0\}$ with maximal margin can be given by

$$\begin{array}{lll} \text{(GP)} & \text{Maximize} & \eta \\ & \text{subject to} & A\boldsymbol{w} + b\mathbf{1} \geqq \eta\mathbf{1} \\ & & B\boldsymbol{w} + b\mathbf{1} \leqq -\eta\mathbf{1} \\ & & ||\boldsymbol{w}|| = 1 \end{array}$$

4 Revision of MSM by MOP/GP

One of drawbacks in MSM is the fact that it yields sometimes too complex discrimination boundaries which cause poor genealizaiton ability. In Nakayama-Kagaku (1998), several modifications of MSM are suggested. One of them introduces interior deviations as well as exterior deviations in MSM. This is formulated as a multi-objective programming problem. If only exterior deviations are considered, this is reduced to a goal programming problem, which is the same as the one suggested by Benett-Mangasarian (1992) called RLPD (robust linear programming discrimination). Applying these MOP/GP approaches to MSM, we can obtain smoother discrimination boundary than the original MSM.

Furtheremore, Nakayama-Kagaku (1998) applied a fuzzy programming technique to MSM, because it is more natural to regard the constraints $A^T\boldsymbol{w}+b\mathbf{1} \geqq 0$ and $B^T\boldsymbol{w}+b\mathbf{1} \leqq 0$ as those which are to be satisfied approximately. This approach yields gray zones for discrimination boundaries, in which the data are not decided clearly as of $\mathcal{A}$ or $\mathcal{B}$. However, this is rather natural, because we usually require further investigation on those data as in cases of medical diagnosis.

5 Support Vector Machine

Support vector machine (SVM) is developed by Vapnik *et al.* (1995), and its main features are

1) SVM is based on linear classfiers with maximal margin on the feature space,
2) SVM uses kernel representation preserving inner products on the feature space,
3) SVM provides an evaluation of the generalization ability using VC dimension.

In cases where training data set X is not linearly separable, we map the original data set X to a feature space Z by some nonlinear map ϕ. Increasing the dimension of the feature space, it is expected that the mapped data set is linearly separable. We try to find linear classifiers with maximal margin in the feature space. Instead of maximizing the minimum interior deviation in (GP) stated above, we use the following equivalent formulation with the normalization $\boldsymbol{w}^T\boldsymbol{z}+b=\pm 1$ at points with the minimum interior deviation:

$$\text{(SVM)}\quad \begin{array}{ll} \text{Minimize} & ||\boldsymbol{w}|| \\ \text{such that} & y_i\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{z}_i+b\right)\geq 1, \quad i=1,\dots,l \end{array}$$

where y_i is $+1$ or -1 depending on the class of $\boldsymbol{z}_i$. Several kinds of norm are possible. When $||\boldsymbol{w}||_2$ is used, the problem is reduced to quadratic programming, while the problem with $||\boldsymbol{w}||_1$ or $||\boldsymbol{w}||_\infty$ is reduced to linear programming (see, e.g., Mangasarian (2000)).

Dual problem of (SVM) with $||\boldsymbol{w}||_2$ is

$$\begin{array}{ll} \text{Min} & \displaystyle\sum_{i=1}^{l}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j y_i y_j \phi(\boldsymbol{x_i})^T\phi(\boldsymbol{x_j}) \\ \text{Subject to} & \alpha_i \geq 0\ ,\ (i=1,\dots,l) \\ & \displaystyle\sum_{i=1}^{l}\alpha_i y_i=0 \end{array} \tag{2}$$

Using the kernel function $K(\boldsymbol{x}, \boldsymbol{x}') = \phi(\boldsymbol{x})^T \phi(\boldsymbol{x}')$, the problem (2) can be reformulated as follows:

$$\begin{array}{ll} \text{Min} & : \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(\boldsymbol{x_i}, \boldsymbol{x_j}) \\ \text{Subject to} : & \alpha_i \geq 0 \ , \ (i = 1, \ldots, l) \\ & \sum_{i=1}^{l} \alpha_i y_i = 0 \end{array} \qquad (3)$$

Several kinds of kernel functions are possible: among them, q-polynomial

$$K(\boldsymbol{x}, \boldsymbol{x}') = (\boldsymbol{x}^T \boldsymbol{x}' + 1)^q$$

and Gaussian

$$K(\boldsymbol{x}, \boldsymbol{x}') = \exp(-\frac{||\boldsymbol{x} - \boldsymbol{x}'||^2}{r^2})$$

are most populary used. In applying the Gaussian kernel, it is important to decide the parameter r. The author and his coresearchers have observed through their numerical experiments that the value of r may be effectively determined by the simple estimate modifying the formula given by Haykin (1994) slightly,

$$r = \frac{d_{max}}{\sqrt[l]{lm}}$$

where d_{max} is the maximal distance among the data; n is the dimension of data; l is the number of data.

Unlike MSM, SVM can provide smooth nonlinear discrimination boundaries in the original data space which result in better generalization ability. However, it can be expected that many devices in MSM and MOP/GP approaches to linear classifiers can be applied to SVM.

Hard Margin and Soft Margin

Separating two sets A and B completely is called the hard margin method, which tends to make overlearning. This implies the hard margin method is easily affected by noise. In order to overcome this difficulty, the soft margin method is introduced. The soft margin method allows some slight error which is represented by a slack variable (exterior deviation) ξ_i $(i = 1, \ldots, l)$. Now, we have the following formulation for the soft margin method:

$$\begin{array}{lll} (\text{SVM}_{\text{soft}}) & \text{Minimize} & \boldsymbol{w}^T \boldsymbol{w} + \sum_{i=1}^{l} \xi_i \\ & \text{subject to} & y_i \left(\mathbf{w}^{\mathrm{T}} \mathbf{x}_i + b \right) \geqq 1 - \xi_i \\ & & \xi_i \geqq 0, \ i = 1, \ldots, l \end{array}$$

Using a kernel function in the dual problem yields

$$\begin{array}{ll} \text{Minimize} & \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(\boldsymbol{x_i}, \boldsymbol{x_j}) \\ \text{Subject to} & C \geq \alpha_i \geq 0 \ , \ (i = 1, \ldots, l) \\ & \sum_{i=1}^{l} \alpha_i y_i = 0 \end{array} \tag{4}$$

It can be seen that the idea of soft margin method is the same as the goal programming approach to linear classifiers. Not only exterior deviations but also interior deviations can be considered in SVM. Such MOP/GP approaches to SVM are discussed in the author and his coresearchers paper (Asada-Nakayama (2001),(2002), Yoon-Nakayama-Yun (2002)). In this event, note that each interior deviation represents how far the sample is from the separating hyperplane, but does not imply the exact distance between the sample and the hyperplane itself. This is a little confusing. Putting the normalization that $\boldsymbol{w}^T \boldsymbol{z} + b = \pm 1$ at support vectors (the samples closest to the separating hyperplane), the corresponding interior deviation indicates the distance between the sample and the hyperplane. However, maximizing η_i subject to $y_i(\boldsymbol{w}^T \boldsymbol{z}_i + b) \geqq 1 + \eta_i$ may yiels unbounded solution, because η_i can increase as much as possible as $y_i(\boldsymbol{w}^T \boldsymbol{z}_i + b) = 1$ tends to the separating hyperplane $\boldsymbol{w}^T \boldsymbol{z}_i + b = 0$.

Example

Let $z_1 = (-1, 1)$, $z_2 = (0, 2) \in \mathcal{A}$ and $z_3 = (1, -1)$, $z_4 = (0, -2) \in \mathcal{B}$. Constraint functions of SVM are given by

$$\begin{array}{ll} z_1: & w_1(-1) + w_2(1) + b \geqq 1 \\ z_2: & w_1(0) + w_2(2) + b \geqq 1 \\ z_3: & w_1(1) + w_2(-1) + b \leqq -1 \\ z_4: & w_1(0) + w_2(-2) + b \leqq -1 \end{array} \tag{5}$$

Since it is clear that the optimal hyperplane has $b = 0$, the constraint functions for z_3 and z_4 are identical to those for z_1 and z_2. The feasible region in (w_1, w_2)-plane is given by $w_2 \geqq w_1 + 1$ and $w_2 \geqq 1/2$. Minimizing the objective function of SVM yields the optimal solution $(w_1,\ w_2) = (-1/2,\ 1/2)$ for the QP formaulation. Similarly, we have a solution among the line segment $\{w_2 \geqq w_1 + 1\} \cap \{-1/2 \leqq w_1 \leqq 0\}$ depending on the initial solution for the LP formulation.

Now consider the goal programming formulation with the objective function consisting of ξ and η. Here, it is clear that $\xi = 0$ at the optimal solution. The constraints include η added in the right hand side. Note that the feasible region in this formulation moves to the north-west by increasing η. Maximizing η yields unbounded optimal solution unless any further constraint in $\boldsymbol{w}$ are added. In MOP/GP approach, therefore, some appropriate normality condition must be imposed on $\boldsymbol{w}$ in order to provide a bounded optimal solution. One of such normality conditions is $||\boldsymbol{w}|| = 1$. However, this normality condition makes the problem to be of nonlinear optimization. Note that the SVM formulation with the objective function minimizing $||\boldsymbol{w}||$ can avoid this unboundedness handily.

If we add the term of η in the objective function of SVM, either an unbounded optimal solution or a bounded optimal solution is possible depending on the trade-off ratio between minimizing $||\boldsymbol{w}||$ and maximizing η. Since it is difficult to decide an appropriate value of the trade-off ratio in practice in advance, some kind of normality condition on η should be imposed. This subject is on-going by the author and his coresearchers.

6 Concluding remarks

A brief survey of linear classifiers using mathematical programming was presented in this paper. In particular, SVM was discussed from a veipoint of MOP/GP. It has been observed that MOP/GP techniques can be effectively applied to these classifiers. However, there remain many problems in question, which will be future subjects.

References

1. Asada, T. and Nakayama, H. (2002) SVM using Multi Objective Linear Programming and Goal Programming, presented at Inter'l Conf. on MOP/GP
2. Bennett, K.P. and Mangasarian, O.L. (1992): Robust Linear Programming Discrimination of Two Linearly Inseparable Sets, *Optimization Methods and Software*, **1**, 23-34
3. Cristianini, N. and Shawe-Taylor, J., (2000) *An Introduction to Support Vector Machines and other kernel-based learning methods*, Cambridge University Press
4. Erenguc, S.S. and Koehler, G.J. (1990) Survey of Mathematical Programming Models and Experimental Results for Linear Discriminant Analysis, *Managerial and Decision Economics*, **11**, 215-225
5. Freed, N. and Glover, F. (1981) Simple but Powerful Goal Programming Models for Discriminant Problems, *European J. of Operational Research*, **7**, 44-60
6. Haykin, S. (1994) *Neural Networks: A Comprehensive Foundation*, Macmillan
7. Mangasarian, O.L. (1968): Multisurface Method of Pattern Separation, *IEEE Transact. on Information Theory*, **IT-14**, 801-807
8. Mangasarian, O.L. (2000) Generalized Support Vector Machines, in *Advances in Large Margin Classifiers*, A.Smola, P.Bartlett, B.Shölkopf, and D.Schuurmans (eds.) Mit Press, Cambridge, pp.135-146
9. Nakayama, H. and Kagaku, N. (1998) Pattern Classification by Linear Goal Programming and its Extensions, *J. of Global Opt.*, **12**, pp.111-126, 1998
10. Nakayama, H. and Asada, T. (2001) Support Vector Machines formulated as Multi Objective Linear Programming, *Proc. of ICOTA2001*, **3**, pp.1171-1178
11. Yoon, M., Nakayama, H. and Yun,Y.B. (2002) A Soft Margin Algorithm controlling Tolerance Directly, presented at International Conference on MOP/GP

Computational Investigations Evidencing Multiple Objectives in Portfolio Optimization

Ralph E. Steuer and Yue Qi

Terry College of Business, University of Georgia, Athens, Georgia 30602-6253 USA

Abstract. In this paper we argue for the recognition of criteria beyond risk and return in portfolio theory in finance. We discuss how multiple criteria are logical and demonstrate computational results consistent with the existence of multiple criteria in portfolio selection. With the efficient frontier becoming an efficient surface, the paper considers that what is the modern portfolio theory of today is best interpreted as a projection onto two-space of the real multiple criteria portfolio selection problem in higher dimensional space.

1 Introduction

At the foundation of modern portfolio theory (Elton and Gruber 1995 is a representative reference), there is the famous Markowitz portfolio selection model. Today, with little in the way of differences from when it was introduced (Markowitz 1952), the Markowitz portfolio selection model is described as follows. Assume n securities, a initial sum of money to be invested, the beginning of a holding period, and the end of the holding period. Let $x_1, \dots x_n$ denote the *investment proportion* weights. These are the proportions of the initial sum to be invested at the beginning of the holding period in the n securities. Also, let r_i be the random variable for the percent return realized on security i at the end of the holding period. Then r_p, the random variable for the percent return realized on a portfolio at the end of the holding period, is the payoff and, as a function of the x_i, is given by

$$r_p = \sum_{i=1}^{n} r_i x_i$$

The difficulty is that the realizations of the r_i, $1 \leq i \leq n$, are not known at the beginning of the holding period (i.e., at the time the x_i are to be chosen). However, the r_i random variables are assumed to have known expected values $E\{r_i\}$, variances σ_{ii}, and covariances σ_{ij}. In this way, the expected value of r_p is given by the linear function

$$E\{r_p\} = \sum_{i=1}^{n} E\{r_i\} x_i$$

and the predicted standard deviation of r_p is given by the square root of a quadratic function

$$\sigma\{r_p\} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i \sigma_{ij} x_j}$$

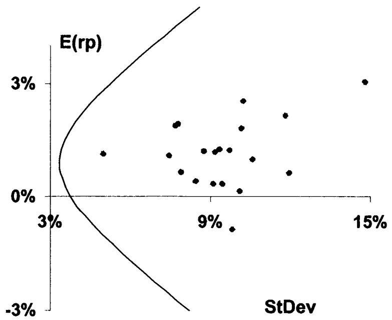
In pursuit of a set of x_i investment proportion weights that will result in a desirable realization of r_p, Markowitz theory assumes that investors will only fix upon both expected portfolio return $E\{r_p\}$ and predicted portfolio standard deviation $\sigma\{r_p\}$ to control the process. Furthermore, Markowitz theory is based upon the assumption that investors like expected portfolio return $E\{r_p\}$, but dislike predicted portfolio standard deviation $\sigma\{r_p\}$. They dislike predicted portfolio standard deviation because standard deviation is believed to capture adequately *risk*, the likelihood that an undesirable realization of r_p might occur. In this way, investors will prefer vectors of investment proportion weights that cause the resultant portfolio to have the smallest predicted standard deviation (i.e., least amount of risk) for any given level of expected return, and investors will prefer vectors of investment proportion weights that cause the resultant portfolio to have the highest expected return for any given level of predicted standard deviation (risk). Thus, the problem is to compute all of the model's *feasible* $(\sigma\{r_p\}, E\{r_p\})$ *nondominated* combinations and then select from them the most preferred. By taking the inverse image of the investor's most preferred nondominated combination, we will then have the x_i investment proportion weights that produce the Markowitz model's "optimal" portfolio.

With regard to the issue of feasibility, $\sum_{i=1}^{n} x_i = 1$ is always a constraint. When this is the only constraint, we have the *short-sales-allowed* model. Because the boundary of the region of all feasible $(\sigma\{r_p\}, E\{r_p\})$ combinations is a hyperbola, the short-sales-allowed model has very nice mathematical characteristics. That is, virtually all information that anyone would ever want to know about the feasible region in $(\sigma\{r_p\}, E\{r_p\})$ space is available in closed-form. One of the places relevant formulas can be found is in (Roll 1977, Appendix A).

For the short-sales-allowed model, the region of all feasible $(\sigma\{r_p\}, E\{r_p\})$ combinations is demonstrated in Figure 1 (left) in which

- **a.** the set of all feasible $(\sigma\{r_p\}, E\{r_p\})$ combinations is unbounded.
- **b.** the dots represent the $(\sigma\{r_i\}, E\{r_i\})$ combinations for the individual securities considered, $1 \leq i \leq n$.
- **c.** the upper half of the hyperbola boundary is the set of all nondominated combinations, referred to in finance as an *efficient frontier.*

A shortcoming of the short-sales-allowed model in which $\sum_{i=1}^{n} x_i = 1$ is the only constraint, is that there is nothing to stop the x_i from taking on negative values. When an x_i is negative, this means that money is raised from



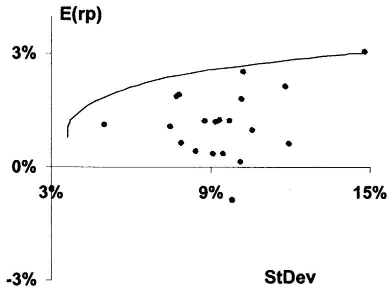


Fig. 1. Short-sales-allowed feasible boundary (left) and short-sales-prohibited efficient frontier (right)

the security. This can only be accomplished by selling *short*. The problem here is that the constraint imposes no limit on the extent to which a stock can be sold short (because the weights can still always be made to sum to one). This is why the feasible region is unbounded. The type of short selling implied by the model is that it is possible to sell a security you don't own to an unlimited extent and use as collateral the stock you are able to buy with the proceeds and the other money you have. While this is possible to a limited extent, carrying it too far will inevitably run into margin requirement difficulties, violate security laws, and not be feasible in reality (despite its theoretical feasibility in the model). Consequently, the region of feasible $(\sigma\{r_p\}, E\{r_p\})$ combinations in the short-sales-allowed model is not nearly as large as commonly portrayed in graphs such as in Figure 1 (left).

In contrast to the short-sales-allowed model, we have the *short-sales-prohibited* model. This model is the same as the short-sales-allowed model but also imposes nonnegativity restrictions on the weights thus "prohibiting" short selling. The inclusion of nonnegativity restrictions might seem innocuous, but they destroy the closed-form solution possibilities of the model and require mathematical programming to be thought of as the solution technique. As a result, a common way to compute the set of all feasible $(\sigma\{r_p\}, E\{r_p\})$ nondominated combinations of the short-sales-prohibited model is to form the mathematical programming problem

$$\begin{aligned} \min \{ \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \sigma_{ij} x_j = \sigma^2\{r_p\}\} \\ \text{s.t.} \quad \sum_{i=1}^{n} E\{r_i\} x_i \geq \rho \\ \sum_{i=1}^{n} x_i = 1 \\ \text{all } x_i \geq 0 \end{aligned}$$

and then solve it repetitively for different values of ρ, that is, for different lower bound values on the expected portfolio return. With a quadratic objective and linear constraints, this is a formulation for which highly effective quadratic/LP solvers are available. In this paper we used, within Excel, the Standard LP/Quadratic solver in Premium Solver Platform V3.5, a Solver upgrade from Frontline Systems (2000). This produces a series of $(\sigma\{r_p\}, E\{r_p\})$ observations which when connected produce the model's "efficient frontier" as in Figure 1 (right). In this case we note that the efficient frontier is not a segment of a hyperbola and that the efficient frontier is not unbounded (having as its rightmost endpoint the security with the highest expected return).

2 Different Perspectives

While anyone with a multiple criteria background would immediately recognize the risk-return portfolio problem as a multiple criteria problem (albeit with only two objectives), mainstream finance does not look at the problem through the same prism and to date has shown no interest in viewing portfolio selection from within a more generalized multiple criteria framework. From their perspective, they feel that they have all that they need. To indicate how portfolio selection is motivated, the following are excerpts from a sampling of top-selling textbooks in finance.

> "The ultimate goal of an investor is an efficient portfolio... Such portfolios aren't necessarily obvious: Investors usually must search out investment alternatives to get the best combinations of risk and return." (Gitman and Joehnk 1999, p. 631)

> "The goal of investors is to maximize wealth. There is a chance that this goal will not be achieved, however, because most investments are risky... To include risk aversion in the decision of security selection, we turn to the *mean-variance criterion.*" (Levy 1999, pp. 193 & 202)

> "In Chapter ... we learned that risky assets should be evaluated on the basis of the expected returns and risk, as measured by the standard deviation... Markowitz portfolio theory provides the way to select optimal portfolios based on using the full information set about securities." (Jones 2000, pp. 511 & 526)

> "Portfolio theory is built around the investor seeking to construct an efficient portfolio that offers the highest return for a given level of risk or the least amount of risk for a given level of return. Of all the possible efficient portfolios, the individual investor selects the portfolio that offers the highest level of satisfaction or utility." (Mayo 2000, p. 163)

> "Even with identical attitudes toward risk, different households and institutions might choose different investment portfolios because of their differing circumstances... These circumstances impose *constraints* on investor choice. Together, objectives and constraints determine appropriate investment policy." (Bodie, Kane and Marcus 2001, p. 131)

The first quote is indicative of the difficulties many books have in separating the portfolio selection problem from single-criterion ways of thinking. The second quote shows the veering off from a multiple objective conceptualization by use of the frequently employed term "mean-variance criterion." The third and fourth quotes are representative of books that more clearly recognize risk and return as distinct criteria, but typically present the material in a rather dogmatic, this is the only way, fashion. The fifth quote is interesting because it recognizes "differing circumstances" but instructs that if present they be taken into account as constraints. Finance is in denial about multiple criteria, but what is more perplexing is that mainstream finance appears to be annoyed by even having to hear about new ideas in portfolio theory. To them, all avenues in portfolio theory have been exhausted years ago and there is nothing new to be found. Consequently, finance has now moved enmasse to other foraging areas such as econometrics-based empirical studies and stochastic asset pricing studies where a person's research future in mainstream finance is more promising. However, portfolio theory is in need of a serious second look as multiple criteria procedures, unbeknownst to mainstream finance, are now abundantly available.

It is the position of this paper that multiple criteria have always been present in portfolio selection and have consistently manifested themselves in the data of financial research, but have only been recognized as such mostly by people who have benefited from also having backgrounds in other fields (for example Ballestero 2000, Chang, Meade, Beasley and Sharaiha 2000, Ehrgott 2003, Hallerbach and Spronk 2000, Hurson and Zopounidis 1994, Jog, Kaliszewski and Michalowski 1999, Konno and Suzuki 1995, and Mansini, Ogryczak and Speranza 2002).

In multiple criteria there is always a line that must be drawn between what is most appropriately modelled as an objective and what is most appropriately modelled as an constraint, but in mainstream finance the line has most likely been drawn too soon. Rather than ignoring, or treating as insignificant, criteria beyond risk and return, an investor could easily face a situation in which his or her optimal portfolio involves an important balance among criteria such as the following six.

$$
\begin{aligned}
&\max \{\text{return}\} \\
&\min \{\text{risk}\} \\
&\max \{\text{dividends}\} \\
&\max \{\text{social responsibility}\} \\
&\min \{\text{number of securities in a portfolio}\} \\
&\min \{\text{short selling}\}
\end{aligned}
$$

These criteria are not unreasonable. Beyond risk and return it is very plausible that an investor might have criterion concerns about dividends (for providing at least a minimal liquidity stream or for corroborating the health of a security), social responsibility (to favor securities involved in environmentally or socially preferable activities), the number of securities in a portfolio (to minimize the time, headache and distraction involved in monitoring and managing a portfolio), and short selling (to avoid problems with the spouse). It is important that these be modelled as objectives so that investor can explore the trade-offs among the concerns before deciding upon the portfolio that provides the greatest preference.

3 Computational Investigations

If it is true that meaningful multiple (that is, beyond two) criteria exist in portfolio selection, then what is presented as risk-return portfolio selection in traditional finance is merely a two-dimensional projection of the real portfolio selection problem in higher dimensional space. What evidence might there be for such a claim? We can start with the concept of the "market portfolio." The market portfolio (Bodie, Kane and Marcus 2001, p. 233) is at the heart of equilibrium theory in portfolio analysis and is the portfolio for which each security is held in proportion to its market value. The market portfolio is supposed to be everyone's optimal portfolio and is to be on the efficient frontier. But in practice it has been found consistently to be deep below the efficient frontier, in fact, so deep below that this cannot be explained by chance variation. Nevertheless, traditional finance has moved onward essentially agreeing not to be bothered by this anomaly that it has never been able to reconcile.

The impact of multiple criteria in the modelling of portfolio selection is that the efficient frontier becomes an efficient *surface.* Thus, if an optimal portfolio is in the middle of the efficient frontier in risk-return finance, than it may not be unreasonable for an optimal portfolio to be in the middle of the efficient surface in multiple criteria finance. We have been computing a long line of experiments that will be reported elsewhere, but will only report on the most dramatic here to make our point.

In a risk-return portfolio problem, let us assume that the feasible region is an ellipse in two-space as in Figure 2. In this case, the efficient frontier is the portion of the periphery of the ellipse in the second quadrant positioned at the center of the ellipse. Correspondingly, in a k-criteria portfolio problem (with objectives beyond risk and return), let us assume that the feasible region is an *ellipsoid* in k-space. In this case, the efficient surface is the portion of the surface of the ellipsoid in an orthant positioned at the center of the ellipsoid. Now let us assume that the market portfolio, which by theory is efficient, is in the middle of the efficient set. If this is the case, then the market portfolio would be at $\mathbf{z}^2$ on the ellipse. However, if (a) there is a third objective, (b) the feasible region is ellipsoidal in three-space, and (c) the market portfolio is in the middle of the efficient surface in R^3, then the market portfolio would *project* onto risk-return space at $\mathbf{z}^3$. Now if (a) there is a fourth objective, (b) the feasible region is ellipsoidal in four-space, and (c) the market portfolio is in the middle of the efficient surface in R^4, then the market portfolio would project onto risk-return space at $\mathbf{z}^4$. With five objectives, then the market portfolio would project onto risk-return space at $\mathbf{z}^5$, and so forth, becoming deeper and deeper.

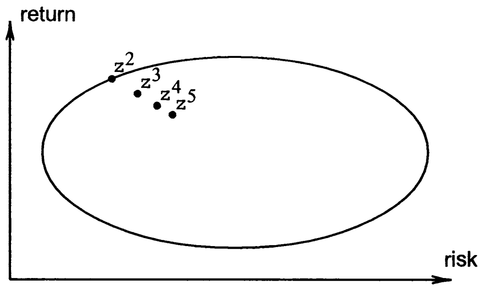


Fig. 2. An ellipsoidal feasible region projected onto two-dimensional risk-return space

To further illustrate, let us look at a 5 objective, 40 constraint, 20 variable multiple objective linear program (MOLP). The problem was created by the random problem generator in ADBASE and then solved for all nondominated vertices in criterion space. Projecting the MOLPs 5,365 nondominated vertices onto the space of any two objectives results in graphs as in Figure 3. Here we can see how a nondominated vertex picked randomly to represent the market portfolio can easily project deep into the interior of the projection of the feasible region onto two-space. Hence, with the empirical results from traditional finance about the "buried" nature of the market portfolio, logical arguments about multiple objectives in portfolio selection, and the projection situation as shown above, there is evidence that the portfolio selection problem would be better modelled within a more generalized multiple criteria framework. Thus, with the efficient frontier becoming an efficient surface, multiple criteria optimization solution procedures as in (Sawaragi, Nakayama and Tanino 1985) would then be more appropriate for searching for optimal portfolios in the new world of multiple criteria portfolio selection.

Fig. 3. Projections of the nondominated vertices of a representative MOLP onto two-space

4 Concluding Remarks

Maybe the reader has noticed that the word optimal in Section 1 was enclosed in quotes. What is optimal depends upon the model. Consider the *literal* model as follows. If an investor's objective is "literally" to maximize the value of r_p at the end of the holding period, then the set of all x_i weighting vector *contenders for optimality* would only include all n unit vectors in R^n. Barring ties, this is because only one of the n securities will have the highest rate of return at the end of the holding period. Thus, to maximize the payoff $r_p = \sum_{i=1}^{n} r_i x_i$ at the end of the holding period, one would only have to have invested in that security alone at the beginning of the holding period. This is in contrast to the Markowitz approach in which we attempt to balance expected portfolio return with predicted portfolio standard deviation, but this comes at a price.

To see this, on the issue of contenders for optimality, let us now discuss the differences between the short-sales-allowed and short-sales-prohibited models and the literal model. In the short-sales-allowed model, the x_i weighting vector contenders for optimality are those that produce the nondominated

$(\sigma\{r_p\}, E\{r_p\})$ combinations along the (unbounded) efficient frontier. However, there are likely no x_i weighting vector contenders for optimality in this model that are in common with the literal model. With regard to the short-sales-prohibited model, there is likely only one x_i weighting vector contender for optimality in this model that is in common with the x_i weighting vector contenders for optimality in the literal model. This is the unit vector in R^n corresponding to the security of highest expected return. The reason for the major differences in the sets of weighting vectors of contenders is that, unfortunately, the literal model's weighting vectors of contenders also contain the weighting vector that minimizes the payoff $r_p = \sum_{i=1}^{n} r_i x_i$ at the end of the holding period. While the Markowitz approach eliminates the possibility of constructing a portfolio that would result in the minimum payoff, its disadvantage is that it eliminates the possibility of maximizing the payoff.

References

1. Ballestero, E. (2000). "Using Compromise Programming in a Stock Market Pricing Model." In Haimes, Y. Y. and R. E. Steuer (eds.), *Research and Practice in Multiple Criteria Decision Making*, Springer, Berlin, 388–399.
2. Bodie, Z., A. Kane and A. J. Marcus (2001). *Essentials of Investments*, 4th edition, McGraw-Hill, Boston.
3. Chang, T.-J., N. Meade, J. E. Beasley and Y. M. Sharaiha (2000). "Heuristics for Cardinally Constrained Portfolio Optimisation," *Computers & Operations Research*, **27**, 1271–1302k.
4. Elton, E. J. and M. J. Gruber (1995). *Modern Portfolio Theory and Investment Analysis*, 5th edition, Wiley, New York.
5. Ehrgott, M., K. Klamroth and C. Schwehm (2003). "An MCDM Approach to Portfolio Optimization," *European Journal Of Operations Research*, forthcoming.
6. Frontline Systems, Inc. (2000). *Premium Solver Platform: User Guide*, Version 3.5, Incline Village, Nevada.
7. Gitman, L. J. and M. D. Joehnk (1999). *Fuindamentals of Investing*, 7th edition, Addison-Wesley, Reading, Massachusetts.
8. Hallerbach, W. and J. Spronk (2000). "A Multi-Dimensional Framework for Portfolio Management." In Karwan, M. H., J. Spronk and J. Wallenius (eds.), *Essays in Decision Making: A Volume in Honour of Stanley Zionts*, Springer, Berlin, 275–293.
9. Hurson, Ch. and C. Zopounidis (1994). "On the Use of Multicriteria Decision Aid Methods to Portfolio Selection." In Climaco, J. (ed.), *Multicriteria Decision Analysis*, Springer, Berlin, 496–507.
10. Jog, V., I. Kaliszewski and W. Michalowski (1999). "Using Attribute Trade-off Information in Investment," *Journal of Multi-Criteria Decision Analysis*, **8**(4), 189–199.
11. Jones, C. P. (2000). *Investments: Analysis and Management*, 7th edition, Wiley, New York.
12. Konno, H. and K.-I. Suzuki (1995). "A Mean-Variance-Skewness Portfolio Optimization Model," *Journal of the Operations Research Society of Japan*, **38**(2), 173–187.

13. Levy, H. (1999). *Introduction to Investments*, 2nd edition, South-Western, Cincinnati.
14. Mayo, H. P. (2000). *Investments: An Introduction*, 6th edition, Harcourt, Fort Worth, Texas.
15. Mansini, R., W. Ogryczak and M. G. Speranza (2002). "LP Solvable Models for Portfolio Selection: A Classification and Computational Comparison," Institute of Control & Computation Engineering, Warsaw University of Technology, Warsaw, Poland.
16. Markowitz, H. (1952). "Portfolio Selection," *Journal of Finance*, **7**(1), 77–91.
17. Roll, R. (1977). "A Critique of the Asset Pricing Theory's Tests," *Journal of Financial Economics*, **4**, 129–176.
18. Steuer, R. E. (2003). "ADBASE: A Multiple Objective Linear Programming Solver for All Efficient Extreme Points and All Unbounded Efficent Edges," Terry College of Business, University of Georgia, Athens, Georgia.
19. Sawaragi, Y., H. Nakayama and T. Tanino (1985). *Theory of Multiobjective Optimization*, Academic Press, Orlando, Florida.

Behavioral Aspects of Decision Analysis with Application to Public Sectors

Hiroyuki Tamura

Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, JAPAN ; tamura@sys.es.osaka-u.ac.jp

Abstract: The focus of this paper is some behavioral (or descriptive) models of individual decision making and group decision making as follows: 1) A model to explain the violations of expected utility models for the individual decision making; 2) A model to describe the ethical consensus formation process among multi-agent conflicting decision makers. Some applications to public sectors are mentioned.

1. Introduction

The expected utility model has been widely used as a normative model of decision analysis under risk for modeling individual decision making. But, various paradoxes [1,2] have been reported for the expected utility model, and it is argued that the expected utility model is not an adequate behavioral (descriptive) model. In this article some behavioral models are shown to resolve expected utility paradoxes. Some realistic applications to public sectors are mentioned.

In multi-attribute utility analysis Keeney and Raiffa's [3] additive/utility independence has been widely used. If we try to deal with consensus formation process under multiple conflicting agents based on the additive/utility independence, we could only model selfish/stubborn agents. To resolve this restriction, we try to model ethical preference [4] of each agent based on the property of convex dependence [5].

2. Behavioral Models to Resolve Expected Utility Paradoxes

In this section a descriptive extension of the expected utility model to account for various paradoxes is shown using the concept of strength of preference.

2.1 Measurable Value Function Under Risk

Let X be a set of all consequences, $x \in X$, and A be a set of all risky alternatives; a risky alternative $\ell \in A$ is written as

$$\ell = (x_1, x_2, \ldots, x_n : p_1, p_2, \ldots, p_n) \tag{1}$$

which yields consequence $x_i \in X$ with probability p_i, $i = 1,2,\ldots,n$ where $\sum p_i = 1$.

Let A^* be a nonempty subset of $A \times A$, and $\succsim$ and $\succsim^*$ be binary relations on A and A^*, respectively. Relation $\succsim$ could also be a binary relation on X. We interpret $\ell_1 \succsim \ell_2 \, (\ell_1, \ell_2 \in A)$ to mean that ℓ_1 is preferred or indifferent to ℓ_2, and $\ell_1\ell_2 \succsim^* \ell_3\ell_4 \, (\ell_1, \ell_2, \ell_3, \ell_4 \in A)$ to mean that the strength of preference for ℓ_1 over ℓ_2 is greater than or equal to the strength of preference for ℓ_3 over ℓ_4.

We postulate that $(A, A^*, \succsim^*)$ takes a positive difference structure that is based on the axioms described by Kranz et al. [6]. The axioms imply that there exists a real-valued function F on A such that for all $\ell_1, \ell_2, \ell_3, \ell_4 \in A$, if $\ell_1 \succsim \ell_2$ and $\ell_3 \succsim \ell_4$, then

$$\ell_1\ell_2 \succsim^* \ell_3\ell_4 \iff F(\ell_1) - F(\ell_2) \geq F(\ell_3) - F(\ell_4). \tag{2}$$

Since F is unique up to a positive linear transformation, it is a cardinal function. It is natural to hold for $\ell_1, \ell_2, \ell_3 \in A$ that

$$\ell_1\ell_3 \succsim^* \ell_2\ell_3 \iff \ell_1 \succsim \ell_2. \tag{3}$$

Then from eqn.(2) we obtain

$$\ell_1 \succsim \ell_2 \iff F(\ell_1) \geq F(\ell_2). \tag{4}$$

Thus, F is a value function on A and, in view of eqn.(2), it is a measurable value function.

We assume that the decision maker will try to maximize the value (or utility) of a risky alternative $\ell \in A$, which is given by the general form as

$$\max_{\ell \in A} F(\ell) = \max_{\ell \in A} \sum_i f(x_i, p_i) \tag{5}$$

where $f(x, p)$ denotes the value (strength of preference) for a consequence x which comes out with probability p. This function is called the *measurable value function under risk*. The main objectives here are to give an appropriate decomposition and interpretation of $f(x, p)$ and to explore its descriptive implications to account for the various paradoxes.

The model eqn.(5) is reduced to the expected utility form by setting

$$f(x, p) = pu(x) \tag{6}$$

when $u(x)$ is regarded as a von Neumann-Morgenstern utility function. The prospect theory of Kahneman and Tversky [7] is obtained by setting

$$f(x, p) = \pi(p)\, v(x) \tag{7}$$

where $\pi(p)$ denotes a weighting function for probability and $v(x)$ a value function for consequence. In this model the value of each consequence is multiplied by a decision weight for probability (not by probability itself).

Extending this Kahneman-Tversky model we obtain a decomposition form

$$f(x,p) = w(p \mid x)v(x) \tag{8}$$

where

$$w(x \mid p) \equiv \frac{f(x,p)}{f(x,1)}, \quad v(x) \equiv v(x \mid 1), \quad v(x \mid p) \equiv \frac{f(x,p)}{f(x^*,p)}, \tag{9}$$

and x^* denotes the best consequence. The expected utility model, eqn.(6), and Kahneman-Tversky model, eqn.(7), are included in our model, eqn.(8), as a special case. Second equation in eqn.(9) implies that $v(x)$ denotes a measurable value function under certainty. Therefore, our model, eqn.(8), also includes Dyer and Sarin's model [8] as a special case. The model eqn.(8) could also be written as

$$f(x,p) = w(p)\, v(x \mid p), \qquad w(p) \equiv w(p \mid x^*). \tag{10}$$

We assume that

$$f(x,0) = 0, \quad \forall x \in X; \quad f(x^R, p) = 0, \quad \forall p \in [0,1] \tag{11}$$

where $x^R \in X$ denotes the reference point (e.g. status quo). The better region on X compared with x^R is called the gain domain and the worse region the loss domain. We also assume that

$$f(x,p) \geq 0 \text{ in the gain domain; } f(x,p) < 0 \text{ in the loss domain.}$$

It will be shown that the conditional weighting function $w(p \mid x)$ describes the strength of preference for probability under the given conditional level of consequence, and $v(x \mid p)$ describes the strength of preference for consequence under the given conditional level of probability.

For interpreting the descriptive model $f(x,p)$ we need to interpret F such that eqn.(2) holds. For all $x_1, x_2, x_3, x_4 \in X$, $\alpha \in [0,1]$, and $y \in X$ such that $x_1 \succsim x_2 \succsim x_3 \succsim x_4$, we consider four alternatives:

$$\ell_1 = (x_1, y; \alpha, 1-\alpha), \quad \ell_2 = (x_2, y; \alpha, 1-\alpha), \tag{12a}$$

$$\ell_3 = (x_3, y; \alpha, 1-\alpha), \quad \ell_4 = (x_4, y; \alpha, 1-\alpha). \tag{12b}$$

In this case we obtain

$$\ell_1\ell_2 \succsim^* \ell_3\ell_4 \iff f(x_1,\alpha) - f(x_2,\alpha) \geq f(x_3,\alpha) - f(x_4,\alpha) \tag{13a}$$

$$\iff v(x_1 \mid \alpha) - v(x_2 \mid \alpha) \geq v(x_3 \mid \alpha) - v(x_4 \mid \alpha) \tag{13b}$$

Therefore, the value function $v(x \mid p)$ defined in eqn.(9) represents the strength of preference for the risky four alternatives in eqn.(12).

On the other hand, for all $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in [0,1]$, $x \in X$ and $x^R \in X$, we consider four alternatives:

$$\ell_1' = (x, x^R; \alpha_1, 1-\alpha_1), \quad \ell_2' = (x, x^R; \alpha_2, 1-\alpha_2), \tag{14a}$$

$$\ell_3' = (x, x^R; \alpha_3, 1-\alpha_3), \quad \ell_4' = (x, x^R; \alpha_4, 1-\alpha_4). \tag{14b}$$

Then we obtain

$$\ell_1'\ell_2' \succsim^* \ell_3'\ell_4' \iff f(x,\alpha_1) - f(x,\alpha_2) \geq f(x,\alpha_3) - f(x,\alpha_4) \tag{15a}$$

$$\iff w(\alpha_1 \mid x) - w(\alpha_2 \mid x) \geq w(\alpha_3 \mid x) - w(\alpha_4 \mid x) \tag{15b}$$

Therefore, the weighting function defined in eqn.(9) represents the strength of preference for the four risky alternatives in eqn.(14).

The above discussions assert that the descriptive model $f(x,p)$ represents the measurable value function under risk to evaluate the consequence $x \in X$ which comes out with probability p.

The Kahneman-Tversky model of eqn.(7) could explain a so-called certainty effect to resolve the Allais paradox [1]. Our descriptive model $f(x,p)$ could also resolve the Allais paradox.

It is well known that the expected utility model is not an appropriate model for modeling extreme events with low probability and high consequence. In [9] it is shown that our descriptive model could resolve such paradox in the application to public sector.

2.2 Measurable Value Function under Uncertainty

In this section we deal with the case where probability of occurence for each event is unknown. When we describe the degree of ignorance and uncertainty by the basic probability of Dempster-Shafer theory [10] the problem is how to represent the value of a set element to construct a measurable value function under uncertainty based on this concept.

In Dempster-Shafer theory of probability let $\mu(A_i)$ be basic probability which could be assigned by any subset A_i of Θ, where Θ denotes a set containing every possible element. The basic probability $\mu(A_i)$ can be regarded as a semimobile probability mass. Let $\Lambda = 2^{\Theta}$ be a set containing every subset of Θ. Then, the basic probability $\mu(A_i)$ is defined on Λ and takes a value contained in [0,1]. When $\mu(A_i) > 0$, A_i is called the focal element or the set element and the following conditions hold:

$$\mu(\phi) = 0, \quad \sum_{A_i \in \Lambda} \mu(A_i) = 1$$

where ϕ denotes an empty set.

Let the value function under uncertainty based on this basic probability be

$$f^*(B,\mu) = w'(\mu)v^*(B \mid \mu) \tag{16}$$

where B denotes a set element, μ denotes the basic probability, w' denotes the weighting function for the basic probability, and v^* denotes the value function with respect to a set element. The set element B is a subset of $\Lambda = 2^{\Theta}$. Equation (16) is an extended version of the value function, eqn.(10), where an element is extended to a set element and the Bayes' probabilty is extended to the Dempster-Shafer basic probability.

For identifying v^*, we need to find the preference relations among set elements, which is not an easy task. If the number of elements contained in the set Θ is getting larger, and the set element B contains considerbale number of element it is not practical to find v^* as a function of B. To cope with this difficulty we could use some appropriate axiom of dominance.

Our descriptive model $f^*(B,\mu)$ could resolve Ellsburg paradox [2] by restricting a set element B to

$$\Omega = \{(m,M) \in \Theta \times \Theta : m \precsim M\}$$

where m and M denote the worst and the best consequence in the set element B, respectively. In this case eqn.(16) is reduced to

$$f^*(\Omega,\mu) = w'(\mu)v^*(\Omega \mid \mu).$$

Incorporating the Dempster-Shafer probability theory in the descriptive model $f^*(\Omega,\mu)$ of a value function under uncertainty, we could model the lack of belief which could not be modelled by Bayes' probability theory. As the result our descriptive model $f^*(\Omega,\mu)$ could resolve the Ellsburg paradox [2].

This descriptive model could be applied to modeling some public sector problems such as cancer risk problems, global environmental problems, etc. under uncertainty in which probability for each event is not known but probability for some set of events is known.

3. Behavioral Models to Resolve Restrictions of Additive/Utility Independence in Consensus Formation Process

Ethical consensus formation process among multi-agent conflicting decision makers is modeled in this section.

3.1 A Group Disutility Function for Multi-agent Decision Making

Let $D_1 \times D_2$ be a two-attribute space of disutility levels and $d_1 \in D_1$, $d_2 \in D_2$ denote the disutility levels of decision maker (DM) 1 and DM2, respectively.

For a given $d_1 \in D_1$ and $d_2 \in D_2$, a group disutility function on $D_1 \times D_2$ space is defined as $g(d_1,d_2)$. Let us assume that d_1^o and d_2^o denote the worst levels of disutility of DM1 and DM2, respectively, and $d_1{}^*$ and $d_2{}^*$ denote the best levels of disutility of DM1 and DM2, respectively. Given an arbitrary $d_2 \in D_2$ a normalized conditional group disutility function (NCGDF) of DM1 is defined as

$$g_1(d_1 \mid d_2) \equiv \frac{g(d_1,d_2) - g(d_1{}^*,d_2)}{g(d_1^o,d_2) - g(d_1{}^*,d_2)} \quad (17)$$

where it is assumed that

$$g\left(d_1^o, d_2\right) > g\left(d_1^*, d_2\right).$$

It is obvious that

$$g_1\left(d_1^o \mid d_2\right) = 1, \quad g_1\left(d_1^* \mid d_2\right) = 0, \tag{18}$$

that is, NCGDF is normalized and is a *single-attribute* group disutility function. Hence it is easily identified.

The NCGDF for DM2, that is, $g_2\left(d_2 \mid d_1\right)$ can also be defined similarly as

$$g_2\left(d_2 \mid d_1\right) \equiv \frac{g\left(d_1 \mid d_2\right) - g\left(d_1, d_2^*\right)}{g\left(d_1 \mid d_2^o\right) - g\left(d_1, d_2^*\right)} \tag{19}$$

The NCGDF $g_1\left(d_1 \mid d_2\right)$ represents DM1's and $g_2\left(d_2 \mid d_1\right)$ represents DM2's subjective preference for the group disutility as a function of his own disutility level, under the condition that the disutility level of the other DM is given.

If NCGDF $g_1\left(d_1 \mid d_2\right)$ does not depend on the conditional level d_2, then attribute D_1 is utility independent [3] of attribute D_2. If attributes D_1 and D_2 are mutually utility independent, the two-attribute disutility function $g\left(d_1, d_2\right)$ can be described as either a multiplicative or additive form [3].

Suppose

$$g_1\left(d_1 \mid d_2\right) \neq g_1\left(d_1 \mid d_2^*\right) \tag{20}$$

for some $d_2 \in D_2$, that is, utility independence does not hold between two attributes D_1 and D_2. In this case we can use a property of *convex dependence* [5] as a natural extension of utility independence.

The property of convex dependence is defined as follows: attribute D_1 is *m*-th order convex dependent on attribute D_2, denoted $D_1\left(\mathrm{CD}_m\right)D_2$, if there exist distinct $d_2^0, d_2^1, \ldots, d_2^m \in D_2$ and real functions $\lambda_0, \lambda_1, \cdots, \lambda_m$ on D_2 such that NCGDF $g_1\left(d_1 \mid d_2\right)$ can be written as

$$g_1\left(d_1 \mid d_2\right) = \sum_{i=0}^{m} \lambda_i\left(d_2\right) g_1\left(d_1 \mid d_2^i\right), \qquad \sum_{i=0}^{m} \lambda_i\left(d_2\right) = 1. \tag{21}$$

for all $d_1 \in D_1$ and $d_2 \in D_2$, where *m* is the smallest non-negative integer for which this relation holds.

This definition says that, if $D_1\left(\mathrm{CD}_m\right)D_2$, then any NCGDF on D_1 can be described as a convex combination of (*m*+1) NCGDFs with different conditional levels where $\lambda_i\left(d_2\right)$ s are not necessarily non-negative. Especially, when $m = 0$ and $D_1\left(\mathrm{CD}_0\right)D_2$, attribute D_1 is utility independent of attribute D_2.

The algorithm for constructing a *two-attribute* group disutility function is as follows [11]:

Step 1. NCGDFs $g_1(d_1 \mid d_2^o)$, $g_1(d_1 \mid d_2^*)$ and $g_1(d_1 \mid d_2^{0.5})$ are assessed, where $d_2^{0.5}$ denotes the intermediate level of attribute D_2 between the worst level d_2^o and the best level d_2^*.

Step 2. If these NCGDFs are almost identical, $D_1(\mathrm{CD}_0)D_2$ holds. Otherwise, go to Step 3.

Step 3. If the convex combination of $g_1(d_1 \mid d_2^o)$ and $g_1(d_1 \mid d_2^*)$ is almost identical with $g_1(d_1 \mid d_2^{0.5})$, $D_1(\mathrm{CD}_1)D_2$ holds. Otherwise, higher order convex dependence holds. Once the order of convex dependence is found, the decomposition form [5] for two-attribute disutility function can be obtained. Single-attribute NCGDFs play a role of basic elements in the two-attribute group disutility function.

Step 4. By assessing the corner values of a group disutility function in two-attribute space, coefficients of linear terms in the two-attribute group disutility function are obtained [6]. As a result a two-attribute group disutility function is obtained.

In modeling multi-agent decision making with conflicting DMs, NCGDF plays the most important role as it can model various patterns of a DM's preference who is self-centered and selfish or flexible and cooperative, and so forth.

3.2 Consensus Formation Modeling for Multi-agent Decision Making

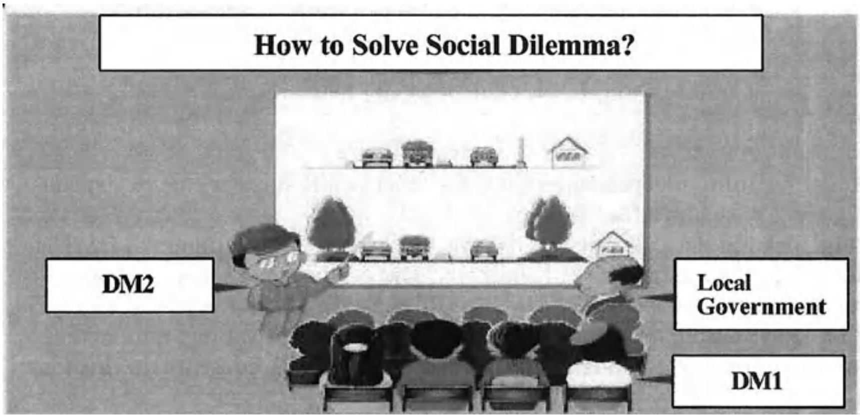


Fig. 1. Consensus formation process between DM1 and DM2.

Let DM1 and DM2 be

DM1: representative of the regional inhabitants;

DM2: representative of the enterpriser who is planning a public project.

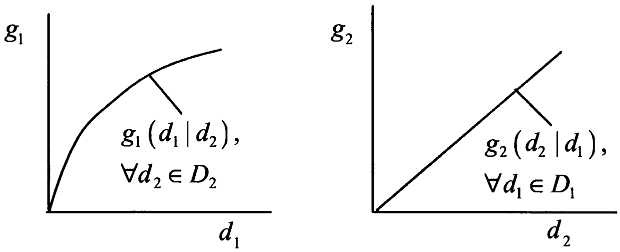
Fig. 1 shows a consensus formation process between DM1 and DM2 where local government plays a role of mediator between them.

Suppose the disutility level d_1 for DM1 evaluates environmental impact from the public project and the disutility level d_2 for DM2 evaluates the cost to realize various countermeasures of the public project. These disutility functions are constructed by questioning the environmental specialists about each situation of DM1 and DM2.

We construct the NCGDFs by again questioning the environmental specialists about each situation of DM1 and DM2. Consequently, suppose we obtained three types of models as follows:

Model 1: Mutual utility independence holds.

Fig. 2 shows the shape of NCGDF for Model 1. Both DM1 and DM2 do not think that group disutility is small unless their own disutility is also small. In this case both DM1 and DM2 are selfish and strongly insist upon their own opinion. This situation shows the initial phase of planning a new project, when the plan has just been presented to the regional inhabitants.



(a) DM1 (b) DM2

Fig. 2. NCGDF of Model 1.

Model 2: Utility independence holds for DM1 and first order convex dependence holds for DM2.

Fig. 3 shows the shape of NCGDF for this Model 2. The attitude of DM1 is almost the same as in Model 1, however, DM2 is becoming more flexible towards obtaining consensus of DM1. In this case DM1 does not have enough information on the project, however, DM2 has obtained various information. This situation corresponds to the second phase of the consensus formation process.

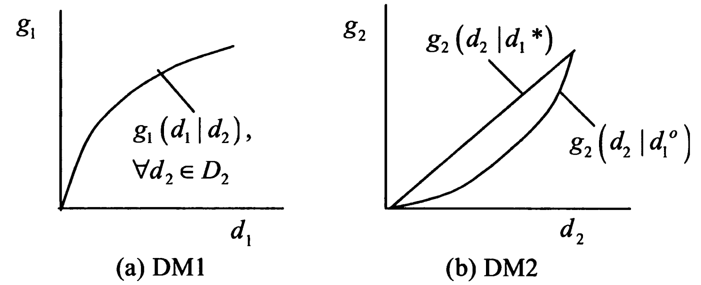


Fig. 3. NCGDF of Model 2.

Model 3: Mutual first order convex dependence holds.

Fig. 4 shows the shape of NCGDF for this Model 3. The attitude of both DM1 and DM2 is getting more flexible and cooperative. In this case both DMs have obtained sufficient information about planning the public project and the countermeasures for preventing environ-mental impacts from the project, and thus, show a mutual concession taking into account ethical consideration with each other. This situation corresponds to the final phase of the consensus formation process between DM1 and DM2.

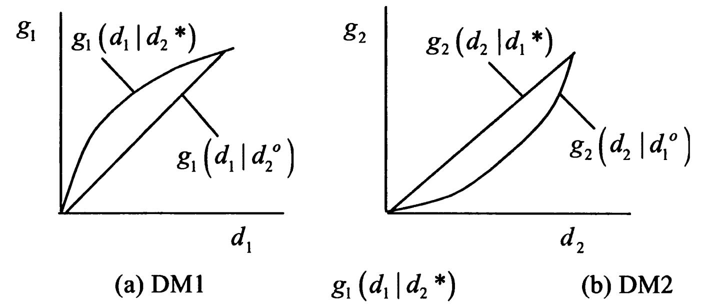


Fig. 4. NCGDF of Model 3.

Suppose the minimum value of group disutility is obtained for Model 3. This implies that the most impartial consensus formation is obtained under the situation of Model 3, which is based on convex dependence between two conflicting DMs.

As seen from the consensus formation model described above it may be used as a fundamental material for discussion when the regional inhabitants and the enterpriser of a public project regulate and adjust their opinion of each other.

4. Concluding Remarks

Behavioral models of decision analysis both in individual decision making and group decision making are described. In the model of the first category consequence dependent non-additive probabilities are introduced as a measurable value function under risk where probability of occurring each event is postulated to be known. The effective application of this approach to public sectors is mentioned in modeling risks of extreme events with low probability and high consequence. Measurable value function under uncertainty is also described where basic probability for a set of events is known but probability of occurring each event is not known. It is shown that Ellsburg paradox is consistently resolved by using this model. Potential applicability of measurable value function under uncertainty to cancer risk problems and global environmental problems is also mentioned.

The model of the second category is described as a group disutility function as a function of normalized conditional group disutility function based on the property of convex dependence among multiple agents. The application of this approach to public sectors is shown in modeling environmental assessment with public participation. It is shown that by using this group disutility model we could model flexible decision makers who could change their attitude on the preference depending upon the disutility level of the other conflicting decision makers. As the result the ethical consensus formation process could be modelled. The consensus formation model described in this paper is expected to be used as a fundamental material for discussion when the enterpriser of a public project and the regional inhabitants regulate and adjust their opinion with each other for realizing better social welfare. A systems method of ethical conflict resolution described in this paper may help to realize a safe, secure and reliable (SSR) megacity.

This research was supported in part by the Japan Society for the Promotion of Science under Grant-in-Aid for Creative Scientific Research (Project No. 13GS0018).

References

1. Allais, M., Hagen, O., eds. (1979) *Expected Utility Hypothesis and the Allais Paradox*. D. Reidel, Dordrecht, Holland.
2. Ellsburg, D. (1961) Risk, ambiguity and the Savage axioms, *Quart. J. Econ.*75, 643-669.
3. Keeney, R.L. and Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Cambridge Univ. Press., UK (originally published by Wiley, New York, 1976)).
4. Sen, A. (1979). *Collective Choice and Social Welfare*, North-Holland, Amsterdam.
5. Tamura, H., and Nakamura, Y. (1983). Decompositions of multi-attribute utility functions based on convex dependence, *Operations Research*, **31**, 488-506.

6. Kranz, D.H., Luce, R.D., Suppes, P. (1971) *Foundations of Measurement*, Academic Press, New York,
7. Kahneman, D., Tversky, A. (1979) Prospect theory: an analysis of decision under risk, Econometrica, **47**, 263-291.
8. Dyer, J.S., Sarin, R.K. (1979) Measurable multi-attribute value function, *Opns. Res.*, **27**, 810-822.
9. Tamura, H., Yamamoto, K., Tomiyama, S., Hatono, I. (2000) Modeling and analysis of decision making problem for mitigating natural disaster risks, *European J. Operational Res.*, **122**, 461-468
10. Shafer, G. (1976) *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ.
11. Tamura, H., and Yoshikawa, T. eds.(1990) *Large-Scale Systems Control and Decision Making*, Marcel Dekker, New York

Optimization Models for Planning Future Energy Systems in Urban Areas

Kiichiro Tsuji

Department of Electrical Engineering,
Graduate School of Engineering, Osaka University
2-1, Yamada-oka, Suita, Osaka 565-0871, Japan

Abstract. A five year JSPS (Japan Society for Promotion of Science) research project titled as "Distributed Autonomous Urban Energy Systems for Mitigating Environmental Impact" was just finished at the end of March 2002. Various types of mathematical optimization models have been developed during the course of the study. Multi-objective optimization models played a major role and many models ended up with mixed integer optimization problems. For solution procedure some of them used commercially available GAMS/Cplex, some of them used decomposition method and the other used heuristic approach such as particle swarm method. These models will be described briefly and also some findings from the project will be shown.

1 Introduction

In Japan energy demand in industry sector has leveled off but in both business and commercial sector and residential sector it is still increasing. The general shift toward tertiary industry and an increase of households have contributed to the increase of energy demand in these sectors. On the other hand, the Kyoto Protocol in 1997 claims 6% reduction in CO2 emission from Japan around the year 2010, and therefore it becomes evermore important to consider efficient use of energy and energy saving in these sectors. As one of the effective means of achieving this, the author proposed the concept of integrated energy service system for specific areas in which electric and thermal energy delivery systems are optimized at the same time. An image of the proposed system is illustrated in Fig.1. Advance of dispersed generation sources and apparatus such as photovoltaic generators, cogeneration for individual buildings, heat pump systems and district heating and cooling systems have created a class of problems to search for the best combination according to a set of objectives.

A five-year JSPS research project titled as "Distributed Autonomous Urban Energy Systems for Mitigating Environmental Impact" was carried out in which the integrated energy service system was explored in detail [1][2]. A number of mathematical programming models have been developed during the course of the project. This paper describes some of these models and some findings from the project will be shown.

2 Optimization Problems in Integrated Energy Service System

In the concept of integrated energy service system, a desirable energy supply system including thermal and electric energy delivery networks is searched for a small specific area (e.g., 2km by 2km). Under these circumstances, the following hierarchical optimization approach has been adopted.

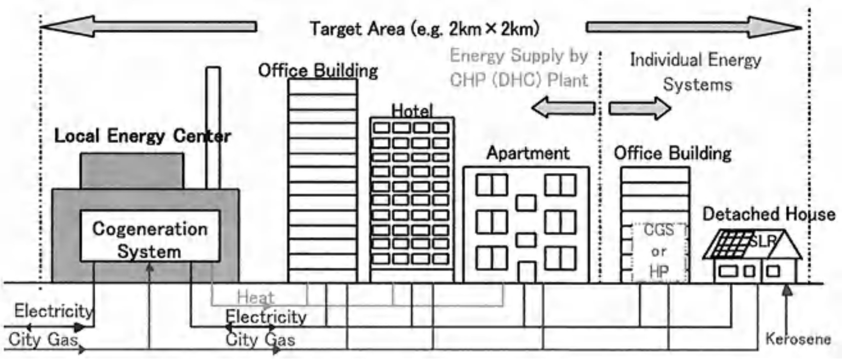


Fig. 1 An image of integrated energy service system

The first problem is to optimize energy systems as a whole for the specific area. The major aim is to look for the best combination of various types of energy systems according to the specific objectives such as CO2 emission from the area, primary energy consumption in the area and cost. CO2 emission versus cost and also primary energy consumption versus cost are two major conflicting pairs of objectives that should be explored by energy system planners forthe area. An important decision output at the end of this optimization is the scale of the area covered by a DHC (District Heating and Cooling) system.

The second problem is an optimization with respect to the DHC system itself. The optimization is carried out over the capacities of each apparatus in the cogeneration plant as well as over network configuration (routing). Here the blocks that are to be connected to the DHC system are predetermined, for example, by the first optimization model above.

The third problem is to optimize electric energy delivery network. In the integrated energy service system, a completely new concept of electric distribution system referred to as FRIENDS (Flexible, Reliable and Intelligent Electric eNergy Delivery System) proposed by Hasegawa and Nara[3] is taken into account. FRIENDS assumes electric power delivery service differentiated by power quality and for this purpose the system includes a number of special facilities called Quality Control Center (QCC). QCC's are connected to each other by high voltage (20kV) distribution line. Customers will be connected to a QCC through low voltage distribution line. On the low voltage side of the QCC, the quality of power will be controlled according to the level of service required. Thus a number of optimization problems will arise that are related to 1) the size of QCC, 2) the location of QCC's, and 3) the distribution network routing.

3 Energy System Optimization for Specific Area

Here the first optimization problem is described in some detail. More concrete description is given in [4].

3.1 Optimization problem

Description of energy systems would be quite complex even for a small specific area because it involves so many factors that must be taken into account. The approach taken here is summarized as follows:

1) The specific area is represented by a number of blocks surrounded by roads and streets.

2) Urban facilities are represented for each block by floor areas for a number of representative buildings category such as office, hotels, retail stores, etc.

3) Energy systems are expressed by a number of alternatives (the structure of each alternative is fixed and is not the object of optimization) for each category of buildings.

4) Energy demand is given for each category of buildings in terms of hourly end-use energy demand (space heating, cooling, heated water supply, cooking, etc.) corresponding to a number of representative days.

5) District heating and cooling (DHC) system is among the energy system alternatives. The decision whether DHC is introduced or not is based on blocks, i.e., each block will determined whether it is to be connected to the DHC thermal energy delivery network. Once a block decides to be connected to the DHC, then every building in the block is supposed to be connected to the DHC network.

6) The optimization over thermal delivery network is based on a simplified assumption that a block is directly connected to the DHC plant by a straight line.

7) The configuration of the DHC co-generation plant is fixed and optimization is carried out only over operational strategies. Energy balance equations and limits on operational variables are the major constraints.

8) The major decision variables for optimization are i) the share of energy system alternatives in terms of floor area and ii) the 0-1 decision variables that represent whether or not a block is connected to the DHC network.

9) Objectives explicitly considered in the model are the cost (fixed and variable costs), CO2 emission and primary energy consumption in the specific area. The multi-objective optimization problem is solved by either weighting method or e-

constraints method. A reference scenario is pre-determined and the set of non-inferior solutions (tradeoff curves) are examined by comparing them with the reference. The information of the tradeoff curves and associated energy system shares is the major outputs of the developed model.

3.2 Issues relevant to optimization

The developed model can be used for various purposes.

1) Tradeoff analysis: The major purpose of the use of the developed model is to analyze the tradeoff between the cost and CO2 emission, or between the cost and primary energy consumption. Figs.2-4 illustrate an example of the analysis. Fig.2 is the tradeoff curve and the variation of energy system configuration in residential sector according to the level of CO2 reduction rate and Fig.3 is the same for the case of primary energy reduction rate. Fig.4 is an image of the specific area under investigation. It was found that the set of Pareto optimal solutions and the associated energy system configurations are different each other. The former case the variation is from electrification system to fuel cell and then to solar energy utilization system, whereas the latter case the DHC system plays an important role.

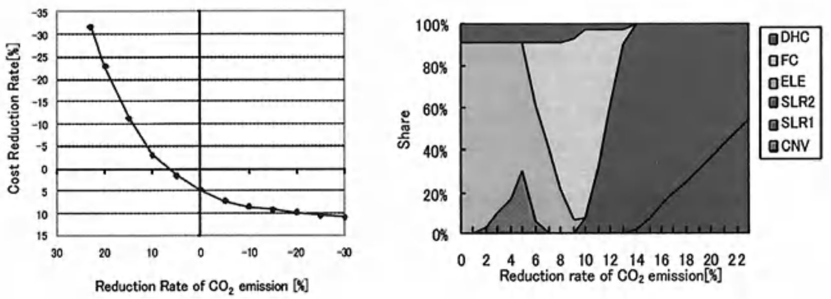


Fig.2 CO2 emission and cost tradeoff, and corresponding energy system share

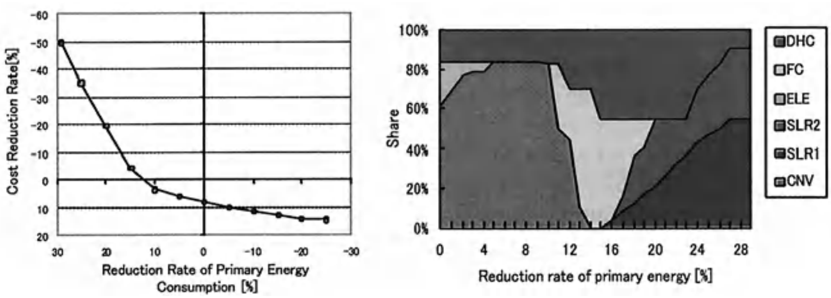


Fig.3 Primary energy and cost tradeoff, and corresponding energy system share

Fig.4 Image of the specific area under study

2) Break even cost analysis for new technologies: There are a number of technologies that are energy efficient but still expensive such as fuel cells, micro GT (Gas Turbine), photovoltaic generation, high COP (Coefficient of Performance) heat pump, solar water heater, wind turbine, etc. The developed model can be used to determine the level of cost that allows a technology to be introduced in the set of non-inferior solutions.

4 Optimization of DHC System[5]

The configuration of the co-generation plant under consideration is shown in Fig.5. This is a gas turbine (GT) co-generation system in which the heat (steam) extracted from GT is used for producing cold water by an absorption refrigerator (RS). If the cold water is insufficient then electric turbo refrigerator (RE) is used to produce required amount of cold water. During wintertime, the heat from GT is used for space heating purposes. Electricity can either be drawn from the utility grid or injected back to the grid. The decision variables in this plant is the capacity of GT in terms of electricity output and thermal output, RS, RE, pump, heat exchanger (HEX) and auxiliary gas boiler (GB).

Fig.6 illustrates the optimization of thermal energy delivery network routes. Here it is assumed that the heat demand is concentrated at the geometrical center of gravity for each block and the thermal energy is delivered only to one of the nodal point for each block. The possible routes are thus predetermined along the streets and roads of blocks. A 0-1 decision variable is allocated to each of the candidate routes expressing whether a pipe should be constructed along the route or not. Constraints are energy balance at each node and limit on thermal transfer for each route. Thermal loss and pressure loss are also taken into account. The cost includes plant equipment, pipes, utility demand and energy. CO2 emission and primary energy consumption can be incorporated.

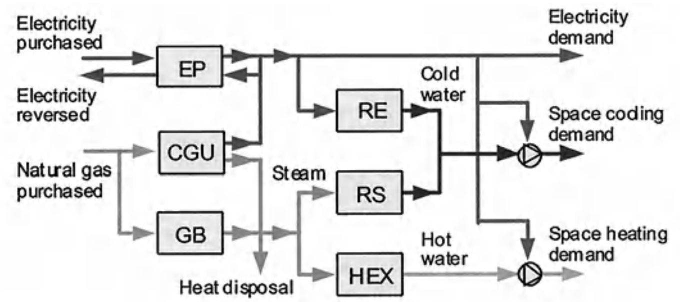


Fig.5 Configuration of cogeneration plant in DHC

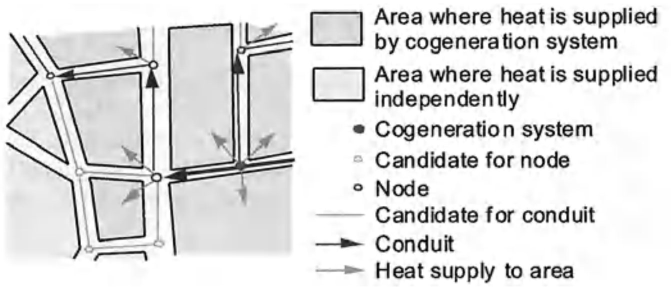


Fig.6 Concept of thermal energy network optimization

The resulting optimization problem is a multi-objective mixed integer programming problem. The size of the problem depends on the scale of the DHC area. The typical size of a block is about 80 meters by 80 meters, therefore if the DHC area is 1 km2 then the number of 0-1 variables will become the order of 100. Moreover, the explicit accounts for thermal and pressure loss makes the problem quite difficult to solve. A decomposition technique is used for solution, but it takes sometimes more than several minutes on a modern fast computer, and tradeoff curves are not easy to obtain. Fig.7 is an example of optimized thermal network.

5 Optimization of Electric Power Distribution Network[6]

Once the DHC area, the size of the DHC plant and the optimal energy system configuration are determined, then electricity demand for each block is completely determined. Under this circumstance, it is possible to consider optimization problem for electricity delivery network. Optimization problems here are twofold. One is the QCC allocation problem and the other is high voltage distribution network design.

The objective of allocating QCC's can be considered in a number of different ways. The approach here uses an evaluation index in which the volume of copper

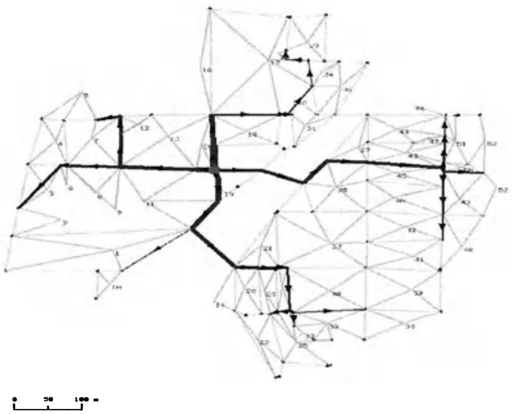


Fig.7 Example of optimized thermal network

required between each QCC and its customers. Voronoi diagram[6] is used to generate initial allocation of the QCC's and then the allocation is modified to minimize the volume of cupper of low voltage distribution line. Different approach may be to determine the location so as to equalize the size of QCC's.

Now the second problem is how to connect each QCC by high voltage distribution lines. Here non-linearity comes in because the circuit equations describing electric power flows for each possible line are non-linear. Any QCC can be equipped with generators and energy storage apparatus such as batteries, so these factors must be taken into consideration as well. Also interruption of electric power must be under prescribed value. The objective function here is taken to be the minimization of distributed generation cost, distribution line construction cost and the loss in the distribution lines. The resulting programming problem is a large scale mixed integer non-linear problem. Conventional approach is not feasible and some heuristic approach must be employed. Here tabu search procedure is incorporated and an optimized solution was found at the expense of several hours of calculation time. Either a better formulation or a better solution procedure is needed and it remains as one of the future research topics. Fig.8 is an example of the results of optimization on QCC location and high voltage distribution network.

6 Concluding Remarks

Optimization problems for urban energy planning appeared in the JSPS research project have been described. Some conclusions are as follows:

1) Some of them are multi-objective and tradeoffs among CO2 emission, primary energy consumption and cost have been discussed.

2) Many problems have been formulated as mixed integer programs: Solution time for large-scale problems is still large.

3) Application to real assessment suggests the importance of real data and also the necessity of computer supported planning systems.

Acknowledgement

The Author is grateful to Prof. Koichi Ito and Ryohei Yokoyama of Osaka Prefecture University, Prof. Koichi Nara of Ibaraki University for their enthusiastic cooperation during the course of the JSPS project.

References

[1] K. Tsuji, Ed., Research Report for 1999, Distributed Autonomous Urban Energy System for Mitigating Environmental Impact, March 2000
[2] Proceedings of International Symposium on "Highly Efficient Use of Energy and Reduction of its Environmental Impact", Osaka International Convention Center, Jan.22-23, 2002
[3] K. Nara, J. Hasegawa, T. Oyama, K. Tsuji and T. Ise (2000), "FRIENDS- fowarding to future power delivery system", Proceedings of 9th IEEE International Conference of Harmonics and Quality of Power, pp.8-18
[4] H. Sugihara and K. Tsuji (2002) Energy -environment-cost tradeoffs in planning energy systems for urban area, Paper presented at MOPGP, Nara
[5] R. Yokoyama and K. Ito (1999) Robust optimal design of a gas turbine cogeneration plant based on minimax regret criterion, 44th ASME Gas Turbine Aeroengine Technical Congress, Exposition and Users Symposium, Paper No.99-GT-128, pp.1-8
[6] K. Nara and Y. Mishima (2002) Optimal network configuration of FRIENDS, in [2]

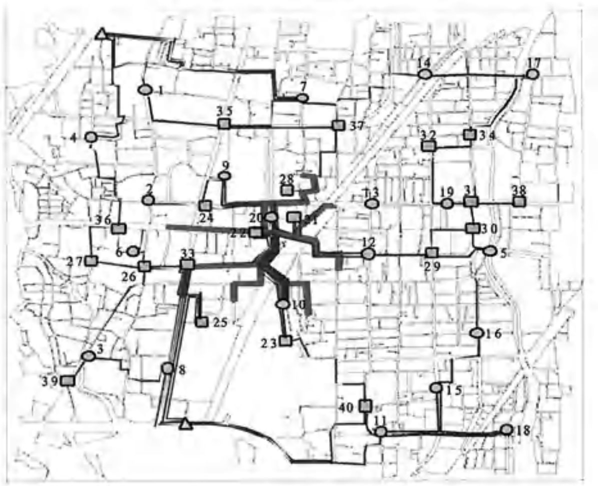


Fig.8 Optimized location of QCC and distribution network

Multiple Objective Decision Making in Past, Present, and Future

Gwo-Hshiung Tzeng

Chairman, Institute of Management of Technology and Environmental Research Group, ;National Chiao Tung University,Taiwan; 1001 Ta-Hsuch Rd., Hsinchu 300, Taiwan; E-mail: ghtzeng@cc.nctu.edu.tw

Abstract

Since Kuhn and Tucker (1951) originally proposed the concept of proper noninferior solution solving nonlinear programming problems and it was later modified by Geoffrion (1967), Yu (1973) further introduce compromise solution method to cope with multicriteria decision-making problems. In addition, Charnes (1955) presented goal programming method, and Bellman and Zadeh (1970) proposed the concepts of decision-making in fuzzy environment, many distinguished work guide person study in this field. This paper review some methods concerning basic mathematical concepts of models applied on multiple objective decision making problem including fuzzy multiobjective linear programming (FMOLP), fuzzy goal programming (FGP), two-phase method, achievement function, data envelopment analysis(DEA), and De Novo Programming.

1. Introduction

Since Kuhn and Tucker (1951) published multiple objectives using vector optimization concept, and Yu (1973) proposed compromise solution method to cope with multicriteria decision-making problems, there have abundant work for applications such as in transportation investment and planning, econometric and development planning, financial planning, business conducting and investment portfolio selecting, land-use planning, water resource management, public policy and environmental issues, and so on. After Bellman and Zadeh (1970) proposed the concepts of decision-making in fuzzy environment, many distinguished work guide person study in this field such as Hwang and Yoon (1981), Zimmermann (1978), Sakawa (1983; 1984a,b), Lee and Li (1993), and so on.

FMOLP formulates the objectives and the constraints as fuzzy sets, characterized by their individual linear membership functions. The decision set is defined as the intersection of all fuzzy sets and the relevant hard constraints. A crisp solution generated by selecting the optimal solution, such that it has the highest degree of membership in the decision set. For further discussions refer to Zimmermann (1978), Werners (1987), Martinson (1993).

This paper organized as follows, the FMOLP model highlighted in Section 2. The FGP model presented in Section 3. The fuzzy goal and fuzzy constraint programming model presented in Section 4. Two phase approach for solving FMOLP problems illustrated in Section 5. Three models of goal programming with achievement function introduced in Section 6. We propose a new multiple objectives programming approach to DEA in Section 7. De Novo programming method in multi-criteria optimal system design presents in Section 8. Finally we summarize most of the methods for multiple objective decision making problems and point out the future direction of our research.

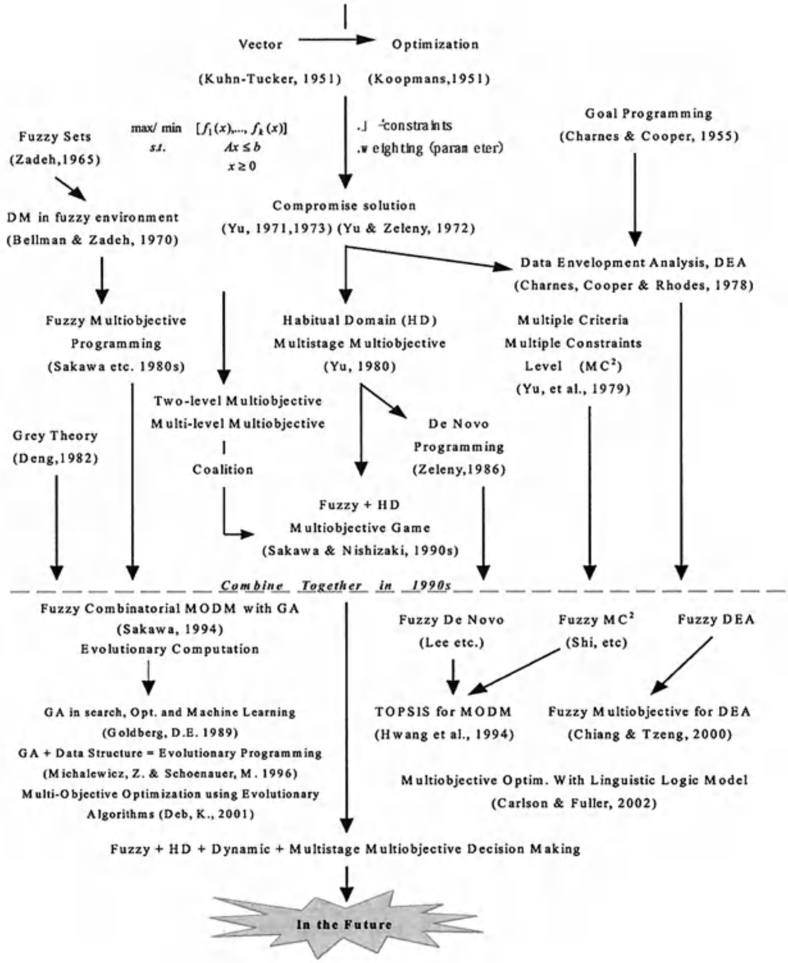


Fig. 1 Development of Multiple Objective Decision Making

2. Fuzzy Multiple Objectives Linear Programming

FMOLP problems usually has the following format:

$$\begin{aligned}
\max\ & \tilde{z}_k = \sum_{j=1}^{n} \tilde{c}_{kj} x_j,\ k = 1,\ 2, \ldots,\ q_1 \\
\min\ & \tilde{w}_k = \sum_{j=1}^{n} \tilde{c}_{kj} x_j,\ k = q_1 + 1, \ldots,\ q \\
s.t.\ & \sum_{j=1}^{n} \tilde{a}_{ij} x_j \le \tilde{b}_i,\ i = 1,\ 2, \ldots,\ m_1;\ \sum_{j=1}^{n} \tilde{a}_{ij} x_j \ge \tilde{b}_i,\ i = m_1 + 1, \ldots,\ m_2 \\
& \sum_{j=1}^{n} \tilde{a}_{ij} x_j = \tilde{b}_i,\ i = m_2 + 1, \ldots,\ m;\ x_j \ge 0,\ j = 1,\ 2, \ldots,\ n
\end{aligned} \tag{1}$$

where $\tilde{c}_{kj}$ is the j-th fuzzy coefficient of the k-th objective, $\tilde{a}_{ij}$ is the j-th fuzzy coefficient of the i-th constraint and $\tilde{b}_i$ is the fuzzy right hand side of the i-th constraint. Problem (1) can solve by transferring it into a crisp model shown as (2).

$$\begin{aligned}
\max\ & (z_k)_\alpha = \sum_{j=1}^{n} (c_{kj})_\alpha^U x_j,\ k = 1,\ 2, \ldots,\ q_1 \\
\min\ & (w_k)_\alpha = \sum_{j=1}^{n} (c_{kj})_\alpha^L x_j,\ k = q_1 + 1,\ \ldots,\ q \\
s.t.\ & \sum_{j=1}^{n} (a_{ij})_\alpha^L x_j \le (b_i)_\alpha^U,\ i = 1,\ 2, \ldots,\ m_1,\ m_2 + 1, \ldots,\ m \\
& \sum_{j=1}^{n} (a_{ij})_\alpha^U x_j \ge (b_i)_\alpha^L,\ i = m_1 + 1,\ \ldots,\ m_2;\ x_j \ge 0, j = 1,\ 2, \ldots,\ n
\end{aligned} \tag{2}$$

where $(c_{kj})_\alpha^U$ and $(c_{kj})_\alpha^L$, $(a_{ij})_\alpha^U$ and $(a_{ij})_\alpha^L$ and $(b_i)_\alpha^U$ and $(b_i)_\alpha^L$ are upper and lower bound of fuzzy number $\tilde{c}_{kj}$, $\tilde{a}_{ij}$ and $\tilde{b}_i$, respectively, by taking α -level cut. Problem (2) can be solved by fuzzy algorithm interactively. For details, see Zimmermann (1978), Lee and Li (1993), Sakawa (1993,1995), Shibano et al. (1996), Shih et al. (1996), Ida and Gen (1997), Shih and Lee (1999) etc.

3. Fuzzy Goal Programming

In most FGP problems can mathematically be represented as:

$$\begin{aligned}
\max\ & [\tilde{f}_1(x), \tilde{f}_2(x), \cdots, \tilde{f}_k(x)] \\
s.t.\ & Ax \le b;\ x \ge 0
\end{aligned} \tag{3}$$

where x, b are vector of variables and right hand side (Yu 1973; Lai et al., 1994) defined the membership function of fuzzy goal as follows:

$$\mu_{g_i}(x) = \begin{cases} 1, & f_i(x) > f_i^*(x) \\ 1 - \dfrac{f_i^*(x) - f_i(x)}{f_i^*(x) - f_i^-(x)}, & f_i^-(x) \le f_i(x) \le f_i^*(x) \\ 0, & f_i(x) < f_i^-(x) \end{cases} \tag{4}$$

where $f_i^*(x)$ and $f_i^-(x)$ represent the positive ideal solution and negative ideal solution, respectively. We can transfer (3) to λ expression method as follows:

$$\begin{aligned} &\max_x \quad \lambda \\ &s.t. \quad \lambda \le \frac{f_i(x) - f_i^-(x)}{f_i^*(x) - f_i^-(x)}, \quad i = 1,\dots,k \\ &\qquad Ax \le b; \; x \ge 0 \end{aligned} \tag{5}$$

We also can employ max-min method to transfer (3) as follows:

$$\begin{aligned} &\max_x \min_i \quad \lambda \\ &s.t. \qquad Ax \le b \\ &\qquad\qquad x \ge 0 \end{aligned} \tag{6}$$

4. Fuzzy Goal and Fuzzy Constraint Programming

The fuzzy goal and fuzzy constraint programming problems can be represented as:

$$\begin{aligned} &\max \quad [\tilde{f}_1(x), \tilde{f}_2(x), \cdots, \tilde{f}_k(x)] \\ &s.t. \quad \tilde{A}x \le \tilde{b} \\ &\qquad x \ge 0 \end{aligned} \tag{7}$$

where x is the vector of variables and $\tilde{b}$ is vector of fuzzy right hand side. First, we define the membership function of fuzzy goal as follows:

$$\mu_{g_i}(x) = \begin{cases} 1, & f_i(x) > f_i^*(x) \\ 1 - \dfrac{f_i^*(x) - f_i(x)}{f_i^*(x) - f_i^-(x)}, & f_i^-(x) \le f_i(x) \le f_i^*(x) \\ 0, & f_i(x) < f_i^-(x) \end{cases} \tag{8}$$

$$\mu_{C_j}(x) = \begin{cases} 1, & (Ax)_j < b_j \\ 1 - \dfrac{(Ax)_j - b_j}{p_j}, & b_j \le (Ax)_j \le b_j + p_j \\ 0, & (Ax)_j > b_j + p_j \end{cases} \tag{9}$$

In this case, we can transfer (7) to λ expression method as follows:

$$\begin{aligned} &\max_x \quad \lambda \\ &s.t. \quad \lambda \le 1 - \frac{f_i(x) - f_i^-(x)}{f_i^*(x) - f_i^-(x)}, \; i = 1,2,\dots,k \\ &\qquad \lambda \le 1 - \frac{(Ax)_j - b_j}{p_j}, \; j = 1,2,\dots,m; \quad x \ge 0 \end{aligned} \tag{10}$$

We also can employ max-min method to transfer (4) as follows:

$$\begin{aligned} &\max_x \min_{i,j} \quad \lambda \\ &s.t. \qquad x \ge 0 \end{aligned} \tag{11}$$

5. Two Phase Approach for Solving FMOLP Problem

Usually there are two or more goals in FMOLP problems, here we illustrate two phase approach for solving the following mathematical programming:

$$\begin{aligned} &\max_x \quad [\tilde{f}_1(\tilde{c}_1,x),\tilde{f}_2(\tilde{c}_2,x),\cdots,\tilde{f}_{k_1}(\tilde{c}_{k_1},x)] \\ &\min_x \quad [\tilde{f}_{k_1+1}(\tilde{c}_{k_1+1},x),\tilde{f}_{k_1+2}(\tilde{c}_{k_1+2},x),\cdots,\tilde{f}_k(\tilde{c}_k,x)] \\ &s.t. \quad \tilde{A}x \Uparrow \tilde{b};\ x\ge 0 \end{aligned} \tag{12}$$

where "$\Uparrow$" represents binary relation and defined as follows:

$\{\Uparrow\}=\{>\}\vee\{\ge\}\vee\{\le\}\vee\{<\}\vee\{=\}$, "$\vee$" means "or".

First, we consider crisp MOLP problems as following programming:

$$\begin{aligned} &\max \quad [\tilde{f}_1(\tilde{C}^U_{1\alpha},x),\tilde{f}_2(\tilde{C}^U_{2\alpha},x),\cdots,\tilde{f}_{k_1}(\tilde{C}^U_{k_1\alpha},x)] \\ &\min \quad [\tilde{f}_{k_1+1}(\tilde{C}^L_{k_1+1,\alpha},x),\tilde{f}_{k_1+2}(\tilde{C}^L_{k_1+2,\alpha},x),\cdots,\tilde{f}_k(\tilde{C}^L_{k\alpha},x)] \\ &s.t. \quad (A)^L_\alpha x \le (b)^U_\alpha \\ &\qquad (A)^U_\alpha x \ge (b)^L_\alpha \\ &\qquad x\ge 0,\quad x\in X_\alpha \end{aligned} \tag{13}$$

Zimmermann (1978) indicated that two important relation between α and β :

(1) Optimal level of α and β , that is $\alpha=\beta$;

(2) Having trade-off relation between α and β .

Then the mathematical programming (13) become as follows:

$$\begin{aligned} &\max_x \quad \beta \\ &s.t. \quad \beta \le \mu_{g_{i(\max)}}(x) \\ &\qquad \beta \ge \mu_{g_{i(\min)}}(x) \\ &\qquad x\in X_\alpha \end{aligned} \tag{14}$$

where

$$\mu_{g_{i(\max)}}(x)=\frac{f_{i(\max)}(C^U_{i\alpha},x)-f^-_{i(\max)\alpha}}{f^*_{i(\max)\alpha}-f^-_{i(\max)\alpha}},\ i=1,2,\ldots,k_1$$

$$\mu_{g_{i(\min)}}(x)=\frac{f^-_{i(\min)\alpha}-f_{i(nin)}(C^L_{i\alpha},x)}{f^-_{i(\min)\alpha}-f^*_{i(\min)\alpha}},\ i=k_1+1,k_1+2,\ldots,k$$

Furthermore, using iteration procedure to find the optimal solution, when $\alpha\cong\beta$, then stop. That is, only to find λ in second phase, such that: $\lambda=\min\{\alpha,\beta\}$

Lee and Li (1993) proposed algorithm for this problems as follows:

Step 1. Setting tolerable error τ , step width ε and initial α -cut ($\alpha=1.0$), iterative frequency $t=1$;

Step 2. Putting $\alpha=\alpha-t\varepsilon$, solve c-LP problem, then obtained β and x;

Step 3. If $|\alpha-\beta|\le\tau$, let $\lambda=\min\{\alpha,\beta\}$, go to step 4; otherwise, go back step 2. If width ε is too large, let $\varepsilon=\varepsilon/2$ and $t=1$, go back step 2;

Step 4. Obtained λ,α,β and x; end.

Therefore, we can solve c-LP2 problems as above two phases algorithm. Moreover, Ida and Gen (1997) proposed following programming to solve this problems:

$$\max \quad \bar{\beta} = \frac{1}{k}\sum_{i=1}^{k} \beta_i$$

$$\begin{aligned} s.t. \quad & \beta \le \beta_i \le \frac{f_{i(\max)}(C_{i\alpha}^U, x) - f_{i(\max)\alpha}^-}{f_{i(\max)\alpha}^* - f_{i(\max)\alpha}^-}, \ i = 1,2,\ldots,k_1 \\ & \beta \le \beta_i \le \frac{f_{i(\min)\alpha}^- - f_{i(nin)}(C_{i\alpha}^L, x)}{f_{i(\min)\alpha}^- - f_{i(\min)\alpha}^*}, \ i = k_1+1, k_1+2,\ldots,k \\ & x \in X_\alpha, \ \beta, \beta_i \in [0,1] \end{aligned} \tag{15}$$

6. Goal Programming with Achievement Functions

Goal programming (GP) is an analytical approach devised to address decision-making problems where targets have been assigned to all the attributes and where the decision-maker is interested in minimizing the non-achievement of the corresponding goals (Romero, 2001).

Initially conceived as an application of single objective linear programming by Charnes and Cooper (1955, 1961), goal programming gained popularity in the 1960s and 70s from the works of Ijiri (1965), Lee (1972), and Ignizio (1976). A key element of a GP model is the achievement function that represents a mathematical expression of the unwanted deviation variables. Each type of achievement function leads to a different GP variant. Tamiz and others (1995) show that around 65% of GP applications reported use lexicographic achievement functions, 21% weighted achievement functions and the rest other types of achievement functions, such as a MINMAX structure in which the maximum deviation is minimized.

The weighted achievement model lists the unwanted deviation variables, each weighted according to importance, the programming shown as (Ignizio 1976):

$$\begin{aligned} Min \quad & \sum_i (\alpha_i d_i^- + \beta_i d_i^+) \\ s.t. \quad & f(x) + d_i^- - d_i^+ = g_i \\ & d_i^- \cdot d_i^+ = 0 \\ & d_i^- \ge 0, \quad d_i^+ \ge 0 \end{aligned} \tag{16}$$

where

$\alpha_i = w_i / k_i$ if d_i^- is unwanted, otherwise $\alpha_i = 0$;

$\beta_i = w_i / k_i$ if d_i^+ is unwanted, otherwise $\beta_i = 0$.

The parameters w_i and k_i are the weights reflecting preferential and normalizing purposes attached to achievement of the i-th goal.

The second model, lexicography achievement model, is made up of an ordered vector whose dimension coincides with the Q number of priority levels established in the model. Each component in this vector represents the unwanted deviation variables of the goals placed in the corresponding priority level (Ignizio 1976).

$$\begin{aligned} Lex \quad Min \quad & a = \left[\sum_{i \in h_1} \left(\alpha_i d_i^- + \beta_i d_i^+ \right), \ldots, \sum_{i \in h_r} \left(\alpha_i d_i^- + \beta_i d_i^+ \right), \ldots, \sum_{i \in h_Q} \left(\alpha_i d_i^- + \beta_i d_i^+ \right) \right] \\ s.t. \quad & f_i(\boldsymbol{x}) + d_i^- - d_i^+ = g_i \quad i \in \{1,\ldots,q\} \quad i \in h_r \quad r \in \{1,\ldots,Q\} \\ & \boldsymbol{x} \in \boldsymbol{F}, \quad d_i^- \geq 0, \quad d_i^+ \geq 0 \end{aligned} \tag{17}$$

where h_r represents the index set of goals placed in the r-th priority level. Lexicographic achievement functions imply a non-compensatory structure of preferences. In other words, there are no finite trade-offs among goals placed in different priority levels (Romero 1991).

The third model, minmax achievement model, seeks for the minimization of the maximum deviation from any single goal. If we represent by D this maximum deviation, the mathematical programming of a LGP model is the following (Flavell 1976):

$$\begin{aligned} \underset{x}{Min} \quad & D \\ s.t. \quad & \alpha_i d_i^- + \beta_i d_i^+ \leq D \\ & f_i(x) + d_i^- - d_i^+ = g_i, \; i \in \{1,\ldots,q\} \\ & x \in F, \; d_i^- \geq 0, \; d_i^+ \geq 0 \end{aligned} \tag{18}$$

This model implies the optimization of a utility function where the maximum deviation is minimized. It provides the most balanced solution among the achievement of different goals. Thus is, it is the solution of maximum equity among the achievement of the different goals(Tamiz and others 1998).

7. Multiple Objective Programming with DEA

Data Envelopment Analysis (DEA) was developed by Charnes, et al. (1978)(CCR model) and extended by Banker et al. (1984)(BCC model), is a non-parametric programming method for estimating production frontiers and evaluating the relative efficiency of decision making units (DMUs), with multiple outputs and multiple inputs. In CCR model, solving the relative efficiency of DMU_k as follows:

$$\begin{aligned} Max \quad & h_k = \sum_{j=1}^{s} u_j y_{jk} \\ s.t. \quad & \sum_{i=1}^{r} v_i x_{ik} = 1, \quad \text{for } k = 1,\ldots,n \\ & \sum_{j=1}^{s} u_j y_{jk} - \sum_{i=1}^{r} v_i x_{ik} \leq 0, \quad \text{for } k = 1,\ldots,n \\ & v_i \geq \varepsilon > 0, \; i = 1,\ldots,r; \quad u_j \geq \varepsilon > 0, \; j = 1,\ldots,s \end{aligned} \tag{19}$$

The objective here is to find the largest sum of weighted outputs of DMU_k while keeping the sum of its weighted inputs at unit value and forcing the ratio of the sum of weighted outputs to the sum of weighted inputs for any DMU to be less than one. Transferring the problem to dual program can then find a minimal value for an intensity factor θ_k that indicates the potential of a proportional reduction in all the inputs of DMU_k.

In BCC model adds another restriction to the envelopment requirements. It requires that the reference point on the production function for DMU_k will be a convex combination of the observed efficient DMUs. The primal formulation for DMU_k is written as:

$$
\begin{aligned}
Max \quad & h_k = \sum_{j=1}^{s} u_j y_{jk} - u_k \\
s.t. \quad & \sum_{i=1}^{r} v_i x_{i0} = 1 \\
& \sum_{j=1}^{s} u_j y_{jk} - \sum_{i=1}^{r} v_i x_{ik} - u_k \le 0, \quad \text{for } k = 1,\dots,n \\
& v_i \ge \varepsilon > 0, i = 1,\dots,r \\
& u_j \ge \varepsilon > 0, j = 1,\dots,s
\end{aligned} \tag{20}
$$

The corresponding primal has a slightly different objective from (19)

Furthermore, considering in CCR model, the efficiency ratio of each DMU is calculated by its own best multipliers, not by the common multipliers for all DMUs. Thus, this model often results in too many DMUs may be identified as efficient. We applied the concept of multiple objectives programming to CCR model to find the common multipliers that could cause the efficiency ratio for all DMU as large as possible. We consider the efficiency ratio of all DMUs rather than k-th DMU_k in CCR model and then establish the following model:

$$
\begin{aligned}
Max \quad & \left[z_1 = \frac{\sum_{r=1}^{s} U_r \cdot y_{r1}}{\sum_{i=1}^{m} V_i \cdot x_{i1}}, z_2 = \frac{\sum_{r=1}^{s} U_r \cdot y_{r2}}{\sum_{i=1}^{m} V_i \cdot x_{i2}}, \dots, z_n = \frac{\sum_{r=1}^{s} U_r \cdot y_{rn}}{\sum_{i=1}^{m} V_i \cdot x_{in}} \right] \\
s.t. \quad & \frac{\sum_{r=1}^{s} U_r \cdot y_{rj}}{\sum_{i=1}^{m} V_i \cdot x_{ij}} \le 1, \ j = 1,\dots,n \\
& U_r \ge \varepsilon > 0, r = 1,\dots,s \\
& V_i \ge \varepsilon > 0, i = 1,\dots,m
\end{aligned} \tag{21}
$$

We further transfer (19) to one objective programming using membership function with fuzzy multiple objectives linear programming approach (Sakawa & Yumine 1983; Sakawa & Yano 1985; Ohta & Yamaguchi 1995), we then conduct the common multipliers to calculate the efficiency achievement for all DMUs, the detail procedure can refer to Chiang & Tzeng(2000).

8. De Novo Programming Method in MODM

Dealing with a MODM problem, we usually confront a situation that is almost impossible to optimize all criteria in a given system. This property is so-called *trade-offs*, which means that one cannot increase the levels of satisfaction for a criterion without decreasing that for another criterion. Zeleny (1981,1986) developed a De Novo programming for designing optimal system by reshaping the feasible set. He suggested that trade-offs are properties of inadequately designed system and thus can be eliminated through designing better, preferably optimal system. Zeleny (1995) proposed the concept of *optimal portfolio of resources* which is design of system resources in the sense of integration, so that there are no trade-offs in a new designed system.

For example, when the budget of designing a new optimal system is higher than total avail budget, Zeleny (1995) suggested an optimum-path ratio to contract the budget to available budget along the optimal path. Along this line, Shi (1995) discussed different budgets from different point of views and define six type optimum-path ratios to find alternatives for optimal system design.

However, since the ideal point used in the De Novo programming is not the ideal point in the ordinary system, the budget for the redesigned system is always larger than the total available budget. Consequently, no matter what optimum-path ratio is used, it only can provide a certain path to locate a solution in the decision space of the new system.

Assuming a MODM problem can be described as follows (Yu, 1985)

$$\begin{aligned} Max \quad & \boldsymbol{Cx} \\ s.t. \quad & \boldsymbol{Ax} \le \boldsymbol{b} \\ & \boldsymbol{x} \ge 0 \end{aligned} \tag{22}$$

where $\boldsymbol{C} = C_{q\times n}$ and $A = A_{m\times n}$, $\boldsymbol{b} = (b_1,...,b_m)^T \in R^m$, and $\boldsymbol{x} = (x_1,...,x_j,...,x_n)^T \in R^n$. Let the kth row of C be denoted by $C^k = (c_1^k,...,c_j^k,...c_n^k) \in R^n$, so that $C^k x, k = 1,...,q$, is the kth criteria or objective function.

Assume that $\boldsymbol{X} = \{\boldsymbol{x} \in R^n | \boldsymbol{Ax} \le \boldsymbol{b}, \boldsymbol{b} \ge 0\}$, the ideal point of (22) is $f^* = (f_1^*,...,f_q^*)^T$, where $f_k^* = \sup\{C^k\boldsymbol{x} | \boldsymbol{x} \in \boldsymbol{X}\}$ for $k = 1,...,q$. If there exists a $x^* = (x_1^*,...,x_n^*)^T \in R^n$, such that $Cx^* = (C^1x^*,...,C^qx^*)^T = (f_1^*,...,f_q^*)^T$, then the x^* called the ideal solution.

Because the components of $\boldsymbol{b}$ in (22) are determined in advance, an ideal point usually is not attainable for the properties of trade-offs among multiple criteria. When the purpose is to design an optimal system rather than optimize a given system, it is of interest to consider following problem:

$$\begin{aligned} Max \quad & \boldsymbol{Cx} \\ s.t. \quad & \boldsymbol{Vx} \le \boldsymbol{B} \\ & \boldsymbol{x} \ge 0 \end{aligned} \tag{23}$$

Then, we find the *Min Vx* for achieving ideal point, i.e.,

$$Min\ Vx$$
$$s.t.\quad C^k x \geq f_k^*,\quad k = 1,\dots,q$$

where $V = pA = (V_1,\dots,V_n) \in R^n$, $p = (p_1,\dots,p_m) \in R^m$ and $B \in R$ present the unit prices of resources and total available budget respectively. Formulation (23) implies that given the unit prices of resources and total available budget allocate the budget, so that the resulting portfolio of resources maximizes the values of the objective functions. There are three methods of De Novo programming for locating a solution while dealing with multi-criteria optimal system design problem: A synthetic-optimal budget, meta-optimal budget, and flexible-constraint meta-optimal budget. For further discussion can refer to Shi (1995).

9. Summary

We have briefly sketched seven important topics of MODM problems, Being space limit, it is difficult to list and discuss many other methods adopted on MODM programming such as fuzzy regression analysis, multiobjective possibilistic/necessity programming, interactive programming methods, two-level/multi-level/multi-stage multiobjective programming, Habitual Domain, Genetic Algorithms and Evolutionary Computing on MODM. We would like to introduce these methods and its applications in near future.

Reference

[1] Banker, R.D., Charnes, A., and Cooper, W.W. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*, **30**(7). 1078-1092.

[2] Bellman, R.E., and Zadeh, L.A. (1970). "Decision-Making in a Fuzzy Environment", *Management Science*, **B17**(1), 141-164.

[3] Charnes, A., and Cooper, W.W. (1961) *Management Models and Industrial Applications of Linear Programming*, John Wiley & Sons, New York.

[4] Charnes, A., Cooper, W.W., and Ferguson, R. (1955). "Optimal Estimation of Executive Compensation by Linear Programming", *Management Science*, **1**(1), 138-151.

[5] Charnes, A., Cooper, W.W., and Rhodes, E. (1978). "Measuring the Efficiency of Decision Making Units" , *European Journal of Operational Research*, **2**(3), 429-444.

[6] Chiang, C.I., and Tzeng, G.H. (2000). "A Multiple Objective Programming Approach to Data Envelopment Analysis", Shi, Y. and Zeleny, M. (eds) *New Frontiers of Decision Making for the Information Technology Era*, World Science Publishing Company, 270-285

[7] Flavell, R.B. (1976). "A New Goal Programming Formulation", *Omega*, **4**(4), 731-732.

[8] Geoffrion, A.M. (1967). "Solving Bicriteria Mathematical Programs", *Operations Research*, **15**(1), 39-54.

[9] Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Springer-Verlag, Addison-Wesley Publishing Company, Inc.

[10] Hwang, C.L., and Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, Heidelberg.

[11] Ida, K., and Gen, M., (1997). "Improvement of Two-Phase Approach for Solving Fuzzy Multiple Objective Linear Programming" *Journal of Japan Society for Fuzzy Theory and System*, **19**(1), 115-121.

[12] Ignizio, J.P. (1976) *Goal Programming and Extensions*, Heath, Lexington, MA.

[13] Ijiri,Y., (1965). *Management Goals and Accounting for Control*, North-Holland, Amsterdam.

[14] Kuhn, H.W., and Tucker, A.W. (1951). "Nonlinear Programming", *Proceedings of the second Berkeley Symposium on Mathematical Statistics and Probability*, 481-492. J. Neyman (Ed.), University of California Press, Berkeley.

[15] Lai, Y.J., Liu, T.Y., and Hwang, C.L. (1994). "TOPSIS for MODM", *European Journal of Operational Research*, **76**(3), 486-500.

[16] Lee, S. (1972). *Goal Programming for Decision Analysis*, Auerbach, PA.

[17] Lee, E.S., and Li, R.J. (1993) "Fuzzy Multiple Objective Programming and Compromise Programming with Pareto Optimum", *Fuzzy Sets and Systems*, **53**(2), 275-288.

[18] Martinson, F.K., (1993). "Fuzzy vs. Min-Max Weighted Multiobjective Linear Programming Illustrative Comparisons," *Decision Sciences*, **24**(5), 809-824.

[19] Ohta, H., and Yamaguchi, T. (1995). "Multi-Goal Programming including Fractional Goal in Consideration of Fuzzy Solutions", *Journal of Japan Society for Fuzzy Theory and System*, **7**(7), 1221-1228.

[20] Romero, C., (1991). *Handbook of Critical Issues in Goal Programming*, Oxford Pergamon Press.

[21] Romero, C. (2001) "Extended Lexicographic Goal Programming: A Unifying Approach" *Omega*, **29**(1), 63-71.

[22] Sakawa, M., (1983). "Interactive Computer Programs for Fuzzy Linear Programming with Multiple Objectives" *International Journal of Man-Machine Studies*, **18**(4), 489-503.

[23] Sakawa, M., (1984). "Interactive Fuzzy Goal Programming for Multiobjective Nonlinear Problems and Its Application to Water Quality Management" *Control and Cybernetics*, **13**(2), 217-228.

[24] Sakawa, M. (1993). *Fuzzy Sets and Interactive Multiobjective Optimization*, Plenum Press, New York.

[25] Sakawa, M., Kato, K., Sundad, H., and Enda, Y., (1995). "An Interactive Fuzzy Satisfying Method for Multiobjective 0-1 Programming Problems through Revised Genetic Algorithms" *Journal of Japan Society for Fuzzy Theory and System*, **17**(2), 361-370.

[26] Sakawa, M. and Yumine, T. (1983). "Interactive Fuzzy Decision Making for Multi-objective Linear Fractional Programming Problems", *Large Scale System*, **5**(1), 105-114.

[27] Sakawa, M. and Yano, H. (1985). "Interactive Decision Making for Multi-objective Linear Fractional Programming Problems with Parameters", *Cybernetics and Systems: An International Journal*, **16**(3), 377-394.

[28] Shi,Y. (1995). "Studies on Optimum-Path Ratios in Multi-Criteria De Novo Programming Problems", *Computers and Mathematics with Applications*, **29**(5), 43-50.

[29] hibano, T., Sakawa, M., and Obata H., (1996). "Interactive Decision Making for Multiobjective 0-1 Programming Problems with Fuzzy Parameters through Genetic Algorithms" *Journal of Japan Society for Fuzzy Theory and System*, **18**(6), 1144-1153.

[30] Shih, H.S., Lai, Y.J., and Lee, E.S. (1996). "Fuzzy Approach for Multi-Level Mathematical Programming Problems" *Computers and Operations Research*, **23**(1), 73-91.

[31] Shih, H.S., and Lee, E.S. (1999). "Fuzzy Multi-Level Minimum Cost Flow Problem" *Fuzzy Sets and Systems*, **107**(2), 159-176.

[32] Tamiz, M., Jones, D.F., and El-Darzi, E. (1995). "A Review of Goal Programming and Its Applications" *Annuals of Operations Research*, **58**(1), 39-53.

[33] Tamiz, M., Jones, D.F., and Romero, C. (1998). "Goal Programming for Decision Making: an Overview of the Current State-of-Art" *European Journal of Operational Research*, **111**(4), 569-581.

[34] Werners, B. (1987). "Interactive Multiple Objective Programming Subject to Flexible Constraints," *European Journal of Operational Research*, **31**(2), 342-349.

[35] Yu, P.L. (1973). "A Class of Solutions for Group Decision Problems", *Management Science*, **19**(8), 936-946.

[36] Yu, P.L. (1985). *Multiple Criteria Decision Making: Concepts, Techniques, and Extensions*, Plenum, New York.

[37] Zeleny, M. (1981). "A Case Study in Multiple Objective Design: De Novo Programming", in *Multiple Criteria Analysis, Operational Methods*, P. Nijkamp and J. Spronk, Eds., Gower Publishing Co., Hampshire, 37-52.

[38] Zeleny, M. (1986). "Optimal System Design with Multiple Criteria: De Novo Programming Approach", *Engineering Costs and Production Economics*, **10**(1), 89-95.

[39] Zeleny, M. (1995). "Trade-Offs-Free Management via De Novo Programming", *International Journal of Operations and Quantitative Management*, **1**(1), 3-13.

[40] Zimmermann, H.J. (1978). "Fuzzy Programming and Linear Programming with Several Objective Functions", *Fuzzy Sets and Systems*, **1**(1), 45-55.

Dynamic Multiple Goals Optimization in Behavior Mechanism

P. L. Yu

C. A. Scupin Distinguished Professor, School of Business, University of Kansas, USA, and Chaired Professor, Institute of Information Management, National Chiao Tung University, Taiwan

C. Y. ChiangLin

Associate Professor, Institute of Finance and Information, National Kaoshiung University of Applied Sciences, Taiwan

Abstract

There exists a behavior mechanism which continuously allocates our attention to various events. Broadly classified, there are seven goals in our life: survival and security, perpetuation of the species, feelings of self-importance, social approval, sensuous gratification, cognitive consistency and curiosity, and self-actualization. For each goal, there is an ideal value or equilibrium point to pursue and/or maintain. If there is a significant discrepancy of current status from the ideal or equilibrium point, a charge (mental pressure) will be produced. At any moment, the totality of all charges created by all goals from all events is called "the charge structure" at that moment. Our mind will try to allocate the attention and resources to reduce the charge to a minimum level. The goals that catch our attention are awakened; otherwise unawakened. The priority for the goals to get our attention follows a dynamic scheme of multiple goals optimization. In this paper, we shall describe and illustrate the dynamic multiple goals optimization, a basic framework of behavior mechanism. Applications to vast decision making problems, especially the challenging ones, will also be mentioned. The mechanism will open up our minds as to make decisions more effectively.

1. Introduction

Human behaviors are undoubtedly dynamic, evolving, interactive and adaptive processes. These complex processes, which evolve dynamic changes of multiple goals, have a common denominator resulting from a common behavior mechanism.

In order to illustrate this dynamic mechanism, let us consider the following example.

Example: First Dating with Motorcycle

John, a college man in Taiwan (where motorcycles are common vehicles), was pretty excited to have his first date at 6:30 PM with his girl friend. At 6 PM, his fantasy of having good time (an important goal), including his girl friend holding him from the back seat, really made him restless. When he arrived at the parking lot for his motorcycle, he was shocked that his motorcycle was locked and he could not unlock it. He called a "taxi" to go to their dating place. When he arrived at the destination he could not find his wallet. The taxi driver was very upset and shouted menacingly, "Shame on you! Young man, there is no free lunch …". John was threatened. How to get out of this situation became his primary concern. Fortunately, he could use cellular phone to call his girl friend to help. John felt extremely embarrassed and his first date was not as excited as he had expected, because he constantly worried about losing his wallet and ID cards therein. After the "date", John went to the parking lot. Surprisingly, he found his motorcycle unlocked and his wallet was still in the compartment under the motorcycle 's seat. Because most motorcycle looks the same, John rationalized that his motorcycle might be locked by mistake by the owner of the motorcycle parked next to his. In any event, what a relief to him!

The above example illustrated the dynamic changes of our goals and behaviors. Indeed our behavior and multi-goal optimization are dynamic, interactive and changing with time and situations. The process can be depicted as in Figure 1. The Figure is self-explanatory. The reader can use the above example and imagination to understand it.

In the next four sections, based on [1-4] we will discuss goal setting and state evaluation, charge structure and attention allocation, least resistance principle and external information input sequentially. Section 6 is for a conclusion.

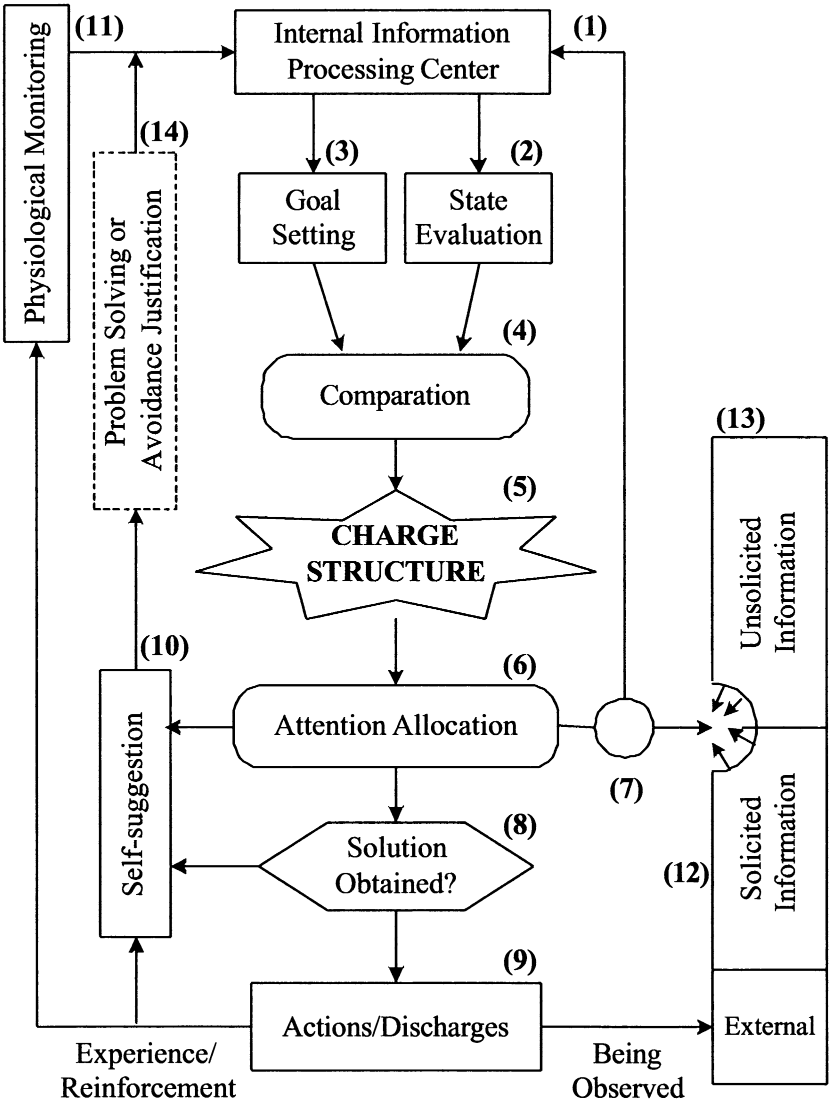


Figure 1. A Behavior Mechanism

2. Goal Setting and State Evaluation

Each human being has a set of goals to reach and maintain. Based on psychology literature, we may summarize these possible goals as listed below:

(1) **Survival and Security:** physiological health (correct blood pressure, body temperature and balance of biochemical states); right level and quality of air, water, food, heat, clothes, shelter and mobility; safety; acquisition of money and other economic goods;
(2) **Perpetuation of the Species:** sexual activities; giving birth to the next generation; family love; health and welfare;
(3) **Feelings of Self-Importance:** self-respect and self-esteem; esteem and respect from others; power and dominance; recognition and prestige; achievement; creativity; superiority; accumulation of money and wealth; giving and accepting sympathy and protectiveness;
(4) **Social Approval:** esteem and respect from others; friendship; affiliation with (desired) groups; conformity with group ideology, beliefs, attitudes and behaviors; giving and accepting sympathy and protectiveness;
(5) **Sensuous Gratification:** sexual; visual; auditory; smell; taste; tactile;
(6) **Cognitive Consistency and Curiosity:** consistency in thinking and opinions; exploring and acquiring knowledge, truth, beauty and religion;
(7) **Self-Actualization:** ability to accept and depend on the self, to cease from identifying with others, to rely on one's own standard, to aspire to the ego-ideal and to detach oneself from social demands and customs when desirable.

The following is a summary of human behavior, called goal setting and state evaluation hypothesis:

> *Each one of us has a set of goal functions and for each goal function we have an ideal state or equilibrium point to reach and maintain (goal setting). We continuously monitor, consciously or subconsciously, where we are relative to the ideal state or equilibrium point (state evaluation). Goal setting and state evaluation are dynamic, interactive and are subject to physiological forces, self-suggestion, external information forces, current data bank (memory) and information processing capacity.*

This hypothesis implies the following points:

(1) There exists a set of goal functions in our internal information processing which are used to measure the many dimensional aspects of life. A probable set is given as above. Goal functions can be mutually associated, interdependent and interrelated.
(2) The goal setting and state evaluation of each goal function are dynamic,

interactive, and subject to physiological forces, self-suggestion, and external information as well as to the current data bank (memory) and information processing capacity.

(3) The influence of self-suggestion can be very pervasive and important to goal setting and state evaluation, and to consequential behavior and decisions. Because of its direct access, self-suggestion can exert its influence on the internal information processing. It can create new perceptions and goal state variables, and can cause restructuring of the data bases in the internal information processing center.

3. Charge Structures and Attention Allocation

Let us first summarize as an important aspect of behavior mechanism as follows:

> *Each event is related to a set of goal functions. When there is an unfavorable deviation of the perceived value from the ideal, each goal function will produce various levels of charge. The totality of the charges by all goal functions is called the **charge structure** and it can change dynamically. At any point in time, our attention will be paid to the event which has the most influence on our charge structure.*

The above is known as charge structure and attention allocation hypothesis. Note that attention allocation is based on dynamic optimization principle. This hypothesis embodies the following details:

(1) Depending on the deviation of the perceived value from the ideal value, various levels of charge for each goal function can occur. The higher level is preemptive over the lower level in obtaining attention.

(2) The collection of the charges on all goal functions created by all current events at one point in time is the charge structure at that moment of time. The charge structure is dynamic and changes (perhaps rapidly) over time. Note, the charge structure can be ordered according to lexicographical ordering to determine its level or strength. For instance, suppose that we have seven goals of concerns as listed above. Two charge structures are given as: A=(5, 2, 5, 4, 3, 2, 1) and B=(2, 3, 5, 5, 4, 2, 1, 3). We could reorder A and B in monotonically decreasing order as A'=(5, 5, 4, 3, 2, 2, 1) and B' = (5, 5, 4, 3, 3, 2, 2, 1). As B' is lexicographical graphically larger than A', B has a higher level of charge structure than A.

(3) Each event can involve many goal functions. Its significance on the charge structure is measured in terms of the extent of which its removal will reduce the levels of charges. Given a fixed set of events, the priority of attention to events at a moment in time depends on the relative significance of the events on the charge structure at that moment in time. The more intense the remaining charge after an event has been removed, the less its relative significance and the lower its relative priority.

(4) For a given set of decision problems with uncertainty, the smaller the decision maker's stake and the greater the decision maker's confidence in obtaining a satisfactory solution within the imposed time limitations, the less the significance of the problems on the charge structures and, consequently, the lower their priority for attention.

4. Least Resistance Principle

In the previous section we described how attention is allotted to various events or decision problems according to the dictates of the charge structure. The event or decision problem with the most significant charge commands our attention at any given moment. Charge structures change quickly as events occur requiring attention to be redirected.

How does our information processing capacity work once attention is allotted to an event? Our information processing function can be characterized by having two modes: (1) **active problem solving** or (2) **avoidance justification.** The former tries to work actively to move the perceived states closer to the ideal states; while the latter tries to rationalize the situations so as to lower the ideal states closer to the perceived states. When operating in either of these modes our information processing function will follow the **least resistance principle** as described below:

> *To release charges, we tend to select the action which yields the lowest remaining charge (the remaining charge is the resistance to the total discharge) and this is called the* ***least resistance principle.***

Note that the least resistance principle is a dynamic optimization principle. The following points should be noted:

(1) Given the charge structure and the set of alternatives at time t, the selected alternative for discharge will be the one that can reduce the residual charge to the lowest level (the least resistance principle).
(2) The majority of daily decision problems that are often repetitive with low stakes and satisfactory solutions are usually readily available for discharge.
(3) When the decision problem involves high stakes and/or uncertainty, active problem solving or avoidance justification can be activated depending on whether or not the decision maker has adequate confidence in finding a satisfactory solution in due time. Either activity can restructure the charge structure and may delay the decision temporarily.
(4) When one is caught unprepared or by surprise (a decision problem involving high stakes and a short time frame for its solution), he/she may act quickly and perhaps unwisely because of time pressures and high levels of charges (refer to the flooded situations of Section 3.2 of [4]).

5. Information Input

Human beings live in a world in which continuous interaction with external events is unavoidable. It can even be suggested that interaction with external events is essential for gaining the information and resources to attain life goals. This interaction with the external world stimulates our information processing mechanism. On the one hand, scanning the external world warns us of forthcoming events that will interfere with our goal attainment. On the other hand, external information is required to confirm or measure our perceived goal states.

We have the following **Information Input Hypothesis**:

> *Humans have innate needs to gather external information. Unless attention is paid, external information inputs may not be processed.*

6. Conclusion

We have briefly described a behavior mechanism and showed that our daily decision problems and behavior are solved or based on a dynamic optimization of a set of goals. Charge structure, attention allocation and arriving information can change the priority of our goals, alternatives and our decisions.

Fortunately, these charge structures, attention allocation and priority over goals can be stabilized and form habitual domains [1-4]. Thus many human behaviors can become predictable. To avoid being trapped by habitual domains. Let us be reminded by Maslow who said, "If the only tool you have is a hammer, you tend to see each problem as a nail."

Let us also face up the challenge that real nontrivial decision problems are dynamic. The optimal solution is a dynamic function of the charge structure, attention allocation and situations of the decision makers.

How to create charges and catch attention by changing situations and sending out information become an important part of forming winning strategies in conflicts, and in solving challenging decision problems. Being limited by space, the interested reader is referred to [2, 4] for further discussions.

References

[1] Chan, S. J. and Yu, P. L., "Stable Habitual Domains: Existence and Implications", Journal of Mathematical Analysis and Applications, Vol. 110, No. 2, pp.469-482, September 1985.

[2] Yu, P. L., Forming Winning Strategies – An Integrated Theory of Habitual Domains, Springer-Verlag, Berlin, Heidelberg, New York, 1990 (392 pages).

[3] Yu, P. L., Habitual Domains, Operations Research, Vol. 39, No. 6, pp. 869-876, 1991

[4] Yu, P. L., Habitual Domains and Forming Winning Strategies, NCTU Press, Hsin-Chu, Taiwan, 2002 (550 pages).

PART II:

General Papers – Theory

An Example-Based Learning Approach to Multi-Objective Programming

Masami Amano and Hiroyuki Okano

IBM Research, Tokyo Research Laboratory 1623-14 Shimotsuruma, Yamato City, Kanagawa 242-8502 (Japan)
E-mail: {amanom,okanoh}@jp.ibm.com

Abstract. In real-world optimization problems involving multiple objectives, the weights of the objectives may not be specified, whereas example solutions, i.e., the solutions for real instances prepared by human experts, are usually available. This paper proposes a method for determining the objective weights by using example solutions as the training set so that a search algorithm can find reasonably good solutions for all the instances. Our proposed method generates neighborhood solutions defined by the search algorithm for each example solution, and determines the weight settings. The method was successfully applied to a scheduling problem in the steel manufacturing industry.

1 Introduction

In real-world optimization problems, system developers often encounter a situation where they can not design the objective function easily. Of course they have the specification, but the constraints specified are so restrictive that most of the real-world instances can not have solutions which satisfy all the constraints. Thus they have to determine how much violation is allowed for each constraint or which constraints are more important than other constraints and design the objective function that takes account of the trade-off between each constraint and each objective.

To determine the trade-off, the most effective method seems to be to ask human experts. However, this might be difficult for various reasons, e.g., the system developers and human experts work for different companies, and so on. To make matters worse, even if we can ask the human experts, they might not be able to evaluate the trade-off quantitatively. As a result, it takes very long time to develop the final systems.

To reduce the development period, there are some clues available, the example solutions for real instances prepared by human experts. Thus what we hope to do is to extract the maximum amount of information from the example solutions and thus design the objective function used in our system. In this paper, we propose a method for determining the weight setting when we fix the objective function in the form of a weighted sum of the objectives, and apply our method to a scheduling problem in the steel manufacturing industry.

However, determining the weight setting by using example solutions has some difficult aspects. For example, human experts are not perfect because the example solutions by human experts are taken as good solutions but it almost cannot happen that the solutions are any of the Pareto optimal solution. Moreover, the evaluation criterion of human experts is not always consistent for all of the example solutions.

Our approach for the problem which we focus on is based on the following conjecture: "If the human expert did the best to find the example solution, the solution is the best solution which he found." When this conjecture is rephrased into the terms of combinatorial optimization, "the example solution is a local optimal solution for the neighborhood structure which the human expert can find," where neighborhood structure means the search space for a solution. Our approach is shown in Section 2 in detail.

We applied our method to a scheduling problem in the steel manufacturing industry [1,2]. In this area, scheduling problems tend to be so complicated and involve conflicting objectives, and many scheduling tasks are still done by skilled human experts. Therefore, it is a challenge for multi-objective programming to emulate the skilled experts.

This paper has the following organization. In Section 2, we show our method to design the objective function (as a weighted sum of the objectives). In Section 3, we give the numerical experiment results by our method and compare the solutions obtained by some variations of our method. In Section 4, we present our conclusions. Due to lack of the space, we omit the detailed explanations and describe them in the full version of this paper [3].

2 Our Learning Approach

In this section, we describe our method that finds the weight settings from example solutions from human experts. Suppose that the objective value of the ith objective is $f_i(s)$ for solution s. The objective function $F(s)$ is defined as

$$F(s) = \sum_i w_i f_i(s),$$

where w_i is the weight for the ith objective. As mentioned before, our approach is based on the idea, "Example solutions are local optimal solutions for the neighborhood structure which human experts can find."

First, we have to define the neighborhood structure of a solution which human experts can find. In this stage, stronger neighborhoods should not be used, because such neighborhoods do not seem to reflect the neighborhoods whose solutions are found by human experts. For example, consider a scheduling problem whose solution is represented as a sequence of jobs sorted in the processed order. The solutions obtained by moving one subsequence to another position will be able to be found by human experts, but the solutions obtained by moving ten subsequences to other positions at a time will not.

Next, we generate inequalities by using the defined neighborhood structure and the example solution. Let s be the example solution and $N(s)$ be the set of the neighborhood solutions of s. Then, if the example solution is better than the neighborhood solutions, the inequality, $F(s) \leq F(s')$, should hold for $s' \in N(s)$. However, this is not always correct because the example solution is not always better than the neighborhood solutions, that is, there may exist a solution s' which dominates solution s. It might be better if, as a preprocess, we change the example solution until the solution becomes the non-dominated solution for the defined neighborhood structure. However, this change may increase the inconsistency between human experts' criterion and the generated inequalities regarding the obtained solution because the changed solution is not the solution made by human experts anymore. We will experiment with both cases, i.e., with preprocessing and without preprocessing, and compare the results. We call both the example solution without preprocessing and the non-dominated solution obtained by preprocessing the *base* solutions.

Finally, we find the weight setting such that they maximize the number of satisfiable inequalities for the set of inequalities generated by the above procedure. This problem can be formalized as a minimization problem for MIP. Let C be a sufficiently large constant, M be the number of inequalities generated, and the neighborhood solution which corresponds to jth inequality be s_j $(1 \leq j \leq M)$. Then this problem is formalized as

$$\begin{array}{ll} \min & \sum_{j=1}^{M} x_j \\ \text{subject to:} & \sum_i (f_i(s) - f_i(s_j)) \cdot w_i \leq C \cdot x_j \\ & \sum_i w_i = 100 \\ & w_i \in \mathbb{R}^+, \quad x_j \in \{0,1\}. \end{array}$$

If the jth inequality is not satisfiable, it is easy to see that $x_j = 1$. Therefore the objective function equals the number of inequalities which are not satisfiable. The second constraint means that the sum of weights is constant. Since the weights are just measures of relative importance, adding the equation does not sacrifice generality. This problem is known to be NP-hard [4]. Therefore it is hard to find the optimal solution unless $P = NP$. We solved this problem by using the IBM OSL [5] and limited the number of nodes for branch-and-bound to obtain a solution within a practical amount of time. Furthermore, to reduce the problem size (the number of inequalities), we delete the neighborhood solutions which are infeasible solutions (the solutions with extremely large violations), or which are dominated by the base solution.

We explained the basic procedure above. The basic procedure is not assumed that we have a priori knowledge that some objectives are more important than the others, but in real situation, we often have it. If so, the objective function is slightly modified and the procedure is also modified. Suppose that the objectives are divided into two sets, A and B, and the objectives in A are more important than the objectives in B. Then the objective function consists of two parts; one consisting of the objectives in A, $F_1(s) = \sum_{i \in A} w_i \cdot f_i(s)$, and the other consisting of the objectives in B, $F_2(s) = \sum_{i \in B} w_i \cdot f_i(s)$. The objective function is defined as

$$F(s) = C \cdot F_1(s) + F_2(s),$$

where C is a sufficiently large constant. Moreover, the procedure to find the weight setting is performed twice. First, the basic procedure is performed for the weights of objectives in A. Second, the basic procedure is performed for the weights of objectives in B by generating inequalities for neighborhood solutions whose objective values for the objectives in A are the same as the base solution's ones. Even if the objectives are divided into three sets or more, its objective function and its learning procedure are performed in the same way.

3 Numerical Experiments

In this section, we present the results obtained by applying our method to a scheduling problem in steel manufacturing industry. The detailed explanation of the scheduling problem is shown in full version of this paper [3].

3.1 Experimental setting

In all there are $m = 8$ objectives. We tried six variations of our methods shown in Table 1. These methods are divided into two types for preprocessing ((2), (4), and (6) in Table 1) or no ((1), (3), and (5)), and divided into three types depending on the degree of a priori knowledge. The methods (1) and (2) do not have any a priori knowledge. The methods (3) and (4) have a priori knowledge that the seven objectives are more important than the other objectives. The methods (5) and (6) have a priori knowledge that the four objectives are the most important, the three are the second, and the one is the least.

Table 1. Six methods examined

(1) no preprocess, no a priori knowledge.
(2) preprocess, no a priori knowledge.
(3) no preprocess, a priori knowledge, seven objs, one obj.
(4) preprocess, a priori knowledge, seven objs, one obj.
(5) no preprocess, a priori knowledge, four objs, three objs, one obj.
(6) preprocess, a priori knowledge, four objs, three objs, one obj.

3.2 Result

We found the weight settings by using the IBM OSL with limiting the number of nodes for branch-and-bound to 20,000 nodes and evaluated the solutions of the problem obtained by using the weight settings found. Since our goal is to find the superior solutions for every objectives, we use the sum of step function as the evaluation function. The evaluation function for solution $s(I)$ is defined as

$$eval(s(I)) = \sum_{i=1}^{m} step(f_i(s(I)), f_i(s_e(I))),$$

where the step function is defined as

$$step(x, y) = \begin{cases} 0, & \text{if } x \leq y \\ 1, & \text{if } x > y, \end{cases}$$

$s_e(I)$ denotes the example solution for instance I, and $f_i(s_e(I))$ denotes the ith objective value of $s_e(I)$. Suppose that $s(I_k)$ is the solution of instance I_k $(1 \leq k \leq K)$ obtained with the weight setting W. Then the evaluation value for the weight setting W is defined as

$$eval(W) = \frac{1}{K} \sum_{k=1}^{K} eval(s(I_k)) \qquad (1)$$

and the evaluation value for the ith objective is defined as

$$eval(obj_i) = \frac{1}{K} \sum_{k=1}^{K} step(f_i(s(I)), f_i(s_e(I))). \qquad (2)$$

This is the procedure we used for evaluating solutions. The number of instances used is 12. We performed each instance ten times for each weight setting. The average results for the solutions are shown in Fig. 1. The data "total" corresponds to the total evaluation value (Eq. (1)), the data "A" corresponds to the sum of the evaluation values for the most important four objectives, and the data "B" corresponds to the sum of the evaluation values for the second important three objectives. From Fig. 1, it is easy to see that preprocessing is useful. We can also see that the total evaluation value of the method (4) is the best. However, this does not strongly support the conclusion that the weight for the method (4) is the best because the superiority of the method (4) appears to come from the better objective values regarding the B. On the other hand, the A's values for the methods (5) and (6) are better than the ones for the method (4), so the B's values in the methods (5) and (6) contribute to the better A's values. Therefore, if a priori knowledge should be respected, the weight setting from the method (5) would be the best weight setting, otherwise the one from the method (4) would be the best one.

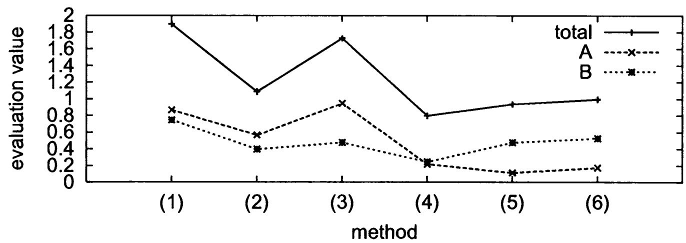


Fig. 1. Evaluation values for weight settings

4 Concluding Remarks

In this paper, we discussed a method to design an objective function by learning from the example solutions from human experts and applied our method to a scheduling problem in the steel manufacturing industry. As a result, we were able to find the good weight settings automatically for the instances used for learning the weight settings.

We offer these observations about our method. First, our method depends on the ability of human experts. If the example solutions from human experts are not good solutions or if the evaluation criterion of the human experts is extremely inconsistent, our method will fail. Second, our method depends on the objective function model. If the objective function does not include the necessary objectives or if the objective function includes irrelevant objectives, such an objective function will not be well fitted to the example solutions. Third, we verified that our weight settings performed very well for the instances used for learning, but not that they perform well for the other instances also. This will be future work.

Finally, the problem we are attacking is very ambiguous and there might not exist a 100% right method, but this is a very common situation in real-world optimization. It seems that this kind of problems should be more focused on for the practical use of optimization.

References

1. P. Cowling. *Optimization in Industry*, chapter on Optimization in steel hot rolling, pages 55–66. John Wiley & Sons, 1995.
2. H. Okano, T. Morioka, and K. Yoda. A Heuristic Solution for the Continuous Galvanizing Line Scheduling Problem in a Steel Mill. IBM Research Report RT0478, 2002.
3. Masami Amano and Hiroyuki Okano. An Example-Based Learning Approach to Multi-Objective Programming. IBM Research Report RT0460, 2002.
4. E. Amaldi and V. Kann. The complexity and approximability of finding maximum feasible subsystems of linear systems. *Theoretical Computer Science*, (147):181–210, 1995.
5. IBM Optimization Solutions and Library. http://www.research.ibm.com/osl/.

Support Vector Machines using Multi Objective Programming

Takeshi Asada[1] and Hirotaka Nakayama[2]

[1] Graduate School of Konan University, Department of Natural Science, 8-9-1, Okamoto, Higashinada-ku, Kobe, 658-8501, Japan
[2] Konan University, Department of Information Science and Systems Engineering, 8-9-1, Okamoto, Higashinada-ku, Kobe, 658-8501, Japan

Abstract. Support Vector Machines are now thought as a powerful method for solving pattern recognition problems. In general, SVMs tend to make overlearning. In order to overcome this difficulty, the notion of soft margin is introduced. In this event, it is difficult to decide the weight for slack variables reflecting soft margin. In this paper, Soft margin method is extended to Multi Objective Linear Programming(MOLP). To solve MOLP, Goal Programming method is used.

1 Principle of SVM

SVMs are usually formulated as Quadratic Programming (QP). However it takes an expensive computation time to solve when the size of data is large. In order to overcome this difficulty, SVMs are reformulated as Linear Programming (LP).

We consider two given sets A and B in n-dimensional real space $\Re^n$. Set $y_i = +1$ for $\mathbf{x}_i \in A$, and set $y_j = -1$ for $\mathbf{x}_j \in B$. When A and B are not linearly separable, the original problem is considered in a feature space usually with a high dimension mapped by some non-linear mapping $\varphi : \mathbf{x} \mapsto \mathbf{z} \quad (\mathbf{x} \in \{\text{original space}\} \ , \ \mathbf{z} \in \{\text{feature space}\})$. Using this mapping, the dataset A and B are expected to be separated linearly. Thus, the separating hyperplane can be expressed [1] as

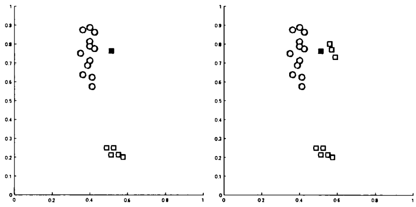
$$\mathbf{w}^{\mathrm{T}}\mathbf{z} + b = 0, \tag{1}$$

and we can lead the Generalized Support Vector Machines [4]:

$$\begin{aligned} \text{[GSVM]} \qquad \text{Minimize} \qquad & ||\mathbf{w}||_q + C\sum_{i=1}^{m}\xi_i \\ \text{s.t.} \qquad & y_i\left(\mathbf{w}^{\mathrm{T}}\mathbf{z}_i + b\right) \geq 1 - \xi_i \qquad (i = 1, \ \ldots, \ m), \end{aligned} \tag{2}$$

where $\mathbf{z}_i = \varphi(\mathbf{x}_i)$, m is the number of data, C is a weight parameter for $\sum \xi_i$, and ξ_i is a slack variable which reflects the distance between separation

[1] When A and B are linearly separable in the dataset, the separating hyperplane can be expressed as $\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$.



(a) (b)

Fig. 1. Example to categorize circles and squares.

hyperplane and misclassified data. Using ℓ_1-norm or ℓ_∞-norm to measure the distance, the problem (2) becomes LP. In this paper, for instance, we use ℓ_∞-norm which yields the following formulation:

$$\begin{aligned} &[\text{LPSVM}] \quad \text{Minimize} \quad ||\mathbf{w}||_1 + C\sum_{i=1}^{m} \xi_i \\ &\qquad\qquad \text{s.t.} \quad y_i\left(\mathbf{w}^{\mathrm{T}}\mathbf{z}_i + b\right) \geq 1 - \xi_i \qquad (i = 1, \ \ldots, \ m) \end{aligned} \tag{3}$$

By the way, the separating hyperplane (1) can be expressed by $f(\mathbf{x}) = \sum_{i=1}^{m} y_i\alpha_i K(\mathbf{x}_i, \mathbf{x}) + b = 0$, where $K(\cdot,\cdot)$ is a kernel function which satisfies Mercer's theorem and defined as $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{z}_i^{\mathrm{T}}\mathbf{z}_j = \varphi(\mathbf{x}_i)^{\mathrm{T}}\varphi(\mathbf{x}_j)$ [4]. In this paper, Gauss function $K(\mathbf{x}, \mathbf{y}) = \exp\left\{-||\mathbf{x} - \mathbf{y}||_2^2/r^2\right\}$ is used as a kernel function. It is important to select an appropriate value of r. From our experience, the formula $r = d_{max}/\sqrt[n]{nm}$ has been observed to give a good estimate. Here, n is the dimension of dataset, m is the number of data, and d_{max} is the maximum distance between the data.

2 Multi Objective Programming formulation

2.1 Maximizing the surplus

Let us consider the example to categorize circles and squares shown in Fig.1 (a). There are 12 circle data and 6 square data. Define the point X as (0.5125, 0.7625) indicated as a black square in Fig.1 (a). The classification results for Fig.1 (a) using formulation (3) is given in Fig.2. Fig.2 (h) is the result by the hard margin method which gives a perfect separation. It can be seen in this example that even though the point X is an outlier, a perfect separation is attained by the hard margin method. On the other hand, the soft margin method (3) which allows some misclassified data can discard the point X as an outlier by controlling the weight of slack variable C. However, it is difficult to select an appropriate value of C. When C is small, soft margin method provides a separating hyperplane which takes into account the influence of noise, shown in Fig.2 (a)-(d). As the value of C increases, a slight change of C brings a big change of separating hyperplanes shown in Fig.2 (d)-(e). This phenomenon implies that soft margin method is sensitive to the value C. To overcome this problem, Multi Objective Programming(MOP) is introduced.

Surplus variable $\boldsymbol{\eta}$ is introduced in addition to slack variable $\boldsymbol{\xi}$. Surplus variable reflects the distance between separating hyperplane and correct data recognized correctly. In order to improve discrimination ability, $\boldsymbol{\eta}$ should be maximized. And in order to control the noise effect, $\boldsymbol{\xi}$ should be minimized. Now the following problem is introduced:

$$\begin{array}{lll} \text{[MOPSVM]} & \text{Minimize} & \sum_{i=1}^{m} \xi_i \\ & \text{Maximize} & \sum_{i=1}^{m} \eta_i \\ & \text{s.t.} & y_i\left(\mathbf{w}^{\mathrm{T}}\mathbf{z}_i + b\right) = 1 - \xi_i + \eta_i \\ & & \boldsymbol{\xi} \geq \mathbf{0}, \quad \boldsymbol{\eta} \geq \mathbf{0} \quad (i = 1, \ \ldots, \ m) \end{array} \tag{4}$$

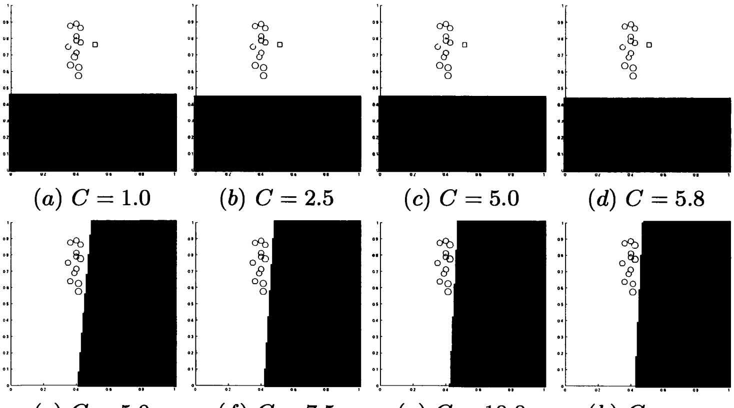
One method for solving MOPSVM is Compromise Programming. Introducing the ideal points ξ^* and η^*, the compromise programming can be expressed as follows [1] [2] [7]:

$$\begin{array}{lll} \text{[MOPSVM-CP]} & \text{Minimize} & ||\mathbf{w}||_1 + d_\xi^+ + d_\eta^- \\ & \text{s.t.} & y_i\left(\mathbf{w}^{\mathrm{T}}\mathbf{z}_i + b\right) = 1 - \xi_i + \eta_i \\ & & \xi^* - \sum_{i=1}^{m} \xi_i + d_\xi^+ - d_\xi^- = 0 \\ & & \eta^* - \sum_{i=1}^{m} \eta_i + d_\eta^+ - d_\eta^- = 0 \\ & & d_\xi^+, d_\xi^-, d_\eta^+, d_\eta^- \geq 0 \\ & & \boldsymbol{\xi} \geq \mathbf{0}, \quad \boldsymbol{\eta} \geq \mathbf{0} \quad (i = 1, \ \ldots, \ m) \end{array} \tag{5}$$

Fig.3 is the result by MOPSVM-CP for the example given in Fig.1 (a). The discrimination boundaries are similarly horizontal separating hyperplanes. However, the area S for the category B (of square data) with fewer data becomes too small depending on the ideal points. This implies that the method may have a poor classification for the category B with fewer number of data. Moreover, in Fig.3, these horizontal separating hyperplanes treat the isolated data X as an outlier. However, if several new square data appear around the point X as the time passes (e.g., Fig.1 (b)), the point X should be considered as an important data rather than an outlier.

2.2 Minimizing the surplus

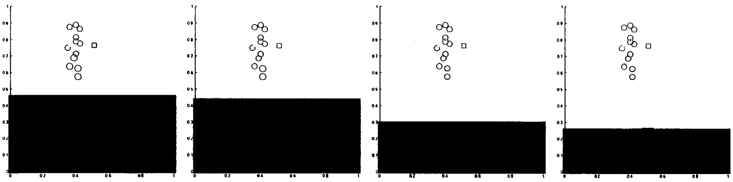
The discriminant boundary provided by MOPSVM-CP is not change suddenly depending on the ideal points, however, the area S becomes too small.



(a) $C = 1.0$ (b) $C = 2.5$ (c) $C = 5.0$ (d) $C = 5.8$

(e) $C = 5.9$ (f) $C = 7.5$ (g) $C = 10.0$ (h) $C \to \infty$

Fig. 2. The result of LPSVM for the example in Fig.1 (*a*).



(a) $\xi^* = 1.0$, $\eta^* = 1.0$ (b) $\xi^* = 1.0$, $\eta^* = 10.0$ (c) $\xi^* = 1.0$, $\eta^* = 100.0$ (d) $\xi^* = 1.0$, $\eta^* = 1000.0$

Fig. 3. The result of MOPSVM-CP for the example in Fig.1 (*a*).

In addition, the point X, which may be important data in the future, is considered as an outlier. In order to overcome the difficulties stated above, $\sum_{i=1}^{m} \eta_i$ is minimized, and the area S becomes larger than the one obtained by MOPSVM-CP. Now the following mathematical programming is introduced:

$$\begin{aligned} &\text{[MOPSVM-pt]} \quad \text{Minimize} \quad && ||\mathbf{w}||_1 + C_1 \sum_{i=1}^{m} \xi_i + C_2 \sum_{i=1}^{m} \eta_i \\ &\text{s.t.} && y_i \left(\mathbf{w}^{\mathrm{T}} \mathbf{z}_i + b \right) = 1 - \xi_i + \eta_i \\ &&& \boldsymbol{\xi} \geq \mathbf{0}, \quad \boldsymbol{\eta} \geq \mathbf{0} \qquad (i = 1, \ \ldots, \ m) \end{aligned} \tag{6}$$

Fig.4 is the result by MOPSVM-pt for the example given in Fig.1 (*a*). This figure shows that MOPSVM-pt makes various separating hyperplanes depend on the parameters C_1 and C_2. This implies that MOPSVM-pt is able to provide the discriminant boundary depending on the environment by controlling C_1 and C_2, e.g., in Fig.4, MOPSVM-pt considers the isolated

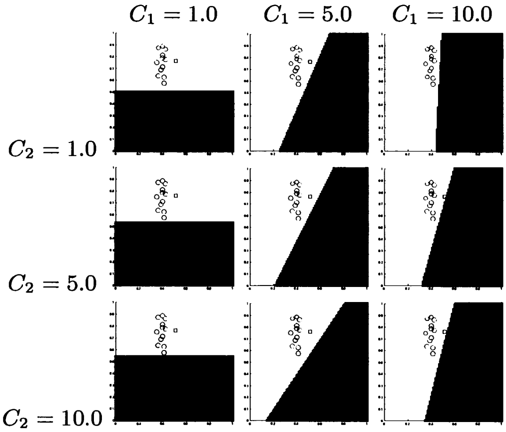


Fig. 4. The result of MOPSVM-pt for the example in Fig.1 (a).

point X as important data when $C_1 = 10.0$, $C_2 = 10.0$, and it considers X as an outlier when $C_1 = 1.0$, $C_2 = 1.0$.

3 Application to Stock Investment problem

In this section, SVM is applied to incremental learning. Incremental learning updates the rule when the knowledge is added. The environment in stock investment tends to change suddenly, therefore incremental learning is expected to yield good results under this situation. The dataset consists of the monthly stock price of some companies from January 1985 to November 1994. It has a seven dimensional economic index. Training data are of the first 49 periods, and test data are of the left 70 periods.

The result using incremental learning is shown in Table 1. In comparison, $C = 10.0$ is used for LPSVM, $\xi^* = 1.0, \eta^* = 10.0$ used for MOPSVM-CP, and $C_1 = 5.0$, $C_2 = 10.0$ used for MOPSVM-pt. Training data "buy" is more than "not to buy", while test data has a reverse situation. Therefore this problem has a difficulty in treating a sudden change of situation. Existing methods tend to overfit to only one category. LPSVM seems to bring better correctness only for the data not to buy, while MOPSVM-CP brings a reverse result. On the other hand, MOPSVM-pt seems to make better correctness for both category.

4 Conclusion

Soft margin method (3) is sensitive to the value C, while the discrimination boundary decided by MOPSVM-CP is not so much sensitive to the ideal

Table 1. Comparison among LPSVM, MOPSVM-CP and MOPSVM-pt in terms of rate of correctness for test data.

		correctness
LPSVM	**Total**	**70.00%**
	to buy	35.29%
	not to buy	81.13%
MOPSVM-CP	**Total**	**33.33%**
	to buy	94.12%
	not to buy	13.46%
MOPSVM-pt	**Total**	**77.14%**
	to buy	35.29%
	not to buy	90.57%

points ξ^* and η^*. However, when the dataset has some unbalance in the number of data, the soft margin method or MOPSVM-CP may give a poor ability of classification for the category with fewer data. On the other hand, MOPSVM-pt seems to work well for each category by controlling the values C_1 and C_2. The decision of C_1 and C_2 from the dataset, i.e., self-tuning of C_1 and C_2 should be subject to a future research.

References

1. Freed N., Glover F. (1981) Simple but powerful goal programming models for discriminant problems. European Journal of Operational Research, Vol.7, 44-60.
2. Erenguc S.S., Koehler G.J. (1990) Survey of Mathematical Programming Models and Experimental Results for Linear Discriminant Analysis. Managerial and Decision Economics, Vol.11, 215-225.
3. Nakayama H., Tanino T. (1994) Theory of Multi Objective Programming and its applications. The Society of Instrument and Control Engineers Edit (in Japanese).
4. Mangasarian O.L. (2000) Generalized Support Vector Machines. In: Smola A., Bartlett P., Schölkopf B., Schuurmans D. (Eds.) Advances in Large Margin Classifiers, Mit Press, Cambridge, 135-146.
5. Bennett K.P., Campbell C. (2000) Support Vector Machines: Hype or Hallelujar?. SIGKDD Explorations, 2, 2.
6. Cristianini N., Shawe-Taylor J. (2000) An Introduction to Support Vector Machines and other Kernel-based learning methods, Cambridge University Press.
7. Shi Y., Peng Y. (2001) Classification for Three-group of Credit Cardholders' Behavior Via A Multiple Criteria Approach. In: Li D. (Ed.) Optimization: Techniques and Applications, 5th International Conference at Hong Kong, December 15-17, 2001. 1279-1286.

On the Decomposition of DEA Inefficiency

Yao Chen, Hiroshi Morita, Joe Zhu[1]

College of Management, University of Massachusetts at Lowell, Lowell, MA 01845 USA
Graduate School of Information Science and Technology, Osaka University, Suita, Osaka, 565-0871, Japan
Department of Management, Worcester Polytechnic Institute, Worcester, MA 01609 USA

Abstract: DEA inefficiency can be characterized by scale and congestion components. We show that the scale and congestion depend on the different technology assumption. The scale component and congestion component depend on the input disposability and returns to scale (RTS) assumptions, respectively. It is also shown that the congestion is associated with the non-zero slack values.

Key words: data envelopment analysis (DEA), efficiency, congestion.

1. Introduction

Byrnes, Färe and Grosskopf (1984) provide a decomposition of Data Envelopment Analysis (DEA) inefficiency into scale and congestion components. McDonald (1996) argues that the decomposition may be sensitive to the order in which the two components are calculated, and consequently, use of the decomposition can result in misleading signals being given to management. However, McDonald (1996) does not notice that different technologies are employed in his two proposed decompositions. As a result, different outcomes of scale and congestion measures should be expected.

By generalizing the definitions of scale and congestion, we show that the scale and congestion component are dependent upon the input disposability and returns to scale (RTS) assumptions without benefit of this insight. McDonald (1996) erroneously concludes that the decomposition is arbitrary. Finally, we characterize congestion in terms of non-zero slack values.

---

[1] Joe Zhu wants to thank the financial support from the Japan Society for Promotion of Science (JSPS). The paper was finished while Joe Zhu was visiting the Osaka University under the JSPS Invitation Research Fellowship.

2. Scale and Congestion Components

On the basis of two different returns to scale (RTS) assumptions and two different input disposability assumptions, one can obtain four different efficiency measures as listed in Table 1.

Table 1. Efficiency measures under different technologies.

Efficiency measures	Returns to Scale (RTS)	Input Disposability
CS	Constant Returns to Scale (CRS)	Strong
CW	Constant Returns to Scale (CRS)	Weak
VS	Variable Returns to Scale (VRS)	Strong
VW	Variable Returns to Scale (VRS)	Weak

Only three of these efficiency measures (CS, VS and VW) were defined and used by Byrnes *et al.* (1984) to define a *specific* scale component and a *specific* congestion component. Within our more general framework a scale component can be obtained by comparing CRS and VRS technologies, and a congestion component can be obtained by comparing strong and weak input disposability technologies. Thus, one could define

Definition 1: (i) (Strong) Scale = CS/VS and (ii) (VRS) Congestion = VS/VW

Obviously, in the above definition, the scale component is based on strong input disposability and the congestion component is based on VRS. On the other hand, one could define:

Definition 2: (i) (Weak) Scale = CW/VW and (ii) (CRS) Congestion = CS/CW where the scale component is based on weak input disposability and the congestion component is based on CRS

Table 2. Three DMUs with two inputs and one output.

	DMU1	DMU2	DMU3
Input 1	1	2	2
Input 2	1	1	2
Output 3	2	4	3

Consider the three DMUs example of McDonald (1996) given in Table 2. Table 3 provides the efficiency results for DMU3 under the different efficiency measures.

From Zhu and Shen (1995) and Seiford and Zhu (1999), DMU3 exhibits strong scale efficiency (CRS) since CS=VS, but weak scale inefficiency (decreasing returns to scale (DRS)) since CW$\neq$VW. Alternate methods for determining scale efficiency (Färe, Grosskopf and Lovell, 1994 and Banker and Thrall, 1992) yield identical results. (Färe, Grosskopf and Lovell (1994) compute

CS/VS =1 and CW/VW < 1, while the Banker and Thrall (1992) approach relies on a CS solution with $\sum \lambda_j^* = 1$ and a unique CW solution with $\sum \lambda_j^* > 1$.)

DEA measure	Intensity vector	Radial component	Input slack
CS	λ_1^* =1.5 or λ_2^* =0.75 or $\lambda_1^* = \lambda_2^*$ =0.5	θ^* =0.75	s_2 =0 s_2 =0.75 s_2 =0.5
CW	λ_1^* =1.5	θ^* =0.75	no slack
VS	$\lambda_1^* = \lambda_2^*$ =0.5	θ^* =0.75	no slack
VW	λ_3^* =1	θ^* =1	no slack

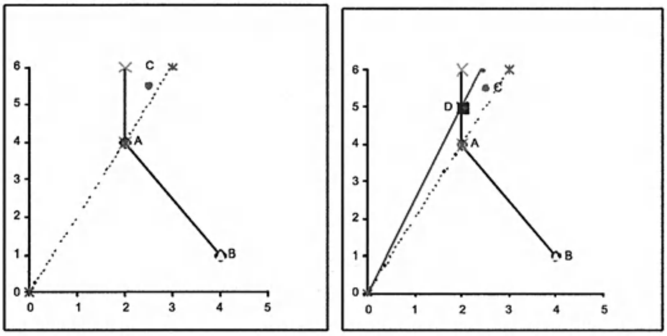


Fig. 1. Congestion at point C and No Congestion at point C

The above RTS estimation demonstrates that different input disposability assumptions may lead to different RTS results for the same DMU. Consequently, the different results for the scale component in definitions 1 and 2 should not be unexpected. In fact, we have (strong) scale =1 and (weak) scale = 0.75. The latter scale inefficiency is due to DRS. i.e., the assumption of weak input disposability causes scale inefficiency. Thus, the scale component is dependent upon the disposability assumption.

Next, note that the two congestion components which McDonald (1996) compared were computed under two different RTS assumptions. In order to further illustrate this point, we first examine the nature of congestion. Figure 1 plots an input isoquant. Input congestion is presented at C in the left Figure 1, but absent at C in the right Figure 1 because of the presence of the weakly efficient point D (a frontier point with non-zero slacks). The shaded regions indicate the presence of input congestion. Because of weak input disposability, the isoquant bends at point A in the left Figure 1 and similarly at point D in the right Figure 1. Furthermore, note that if the efficient reference set consists of A, point C will have a positive slack value for the second input x2. However, for the efficient reference set consisting of points A and D, point C will not have slack values. Thus, we obtain the following theorem:

Theorem: VRS (CRS) input congestion is not presented at DMU_o if and only if there exist some referent frontier DMUs such that non-zero input slack values are not detected for the VRS (CRS) strong disposability DEA measure.

[Proof]: We prove this theorem under definition 1, i.e., VRS. The proof under definition 2, i.e., CRS, is similar.

The VS measure is

$$\begin{aligned}
&\theta^* = \min \theta \\
s.t. \ &\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io} \quad i = 1,\ldots,m; \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} \quad r = 1,\ldots,s; \qquad \text{(VS)} \\
&\sum_{j=1}^{n} \lambda_j = 1; \\
&\lambda_j \ge 0 \ \ and \ \ \theta \ free.
\end{aligned}$$

The VW measure is

$$\begin{aligned}
&\phi^* = \min \phi \\
s.t. \ &\sum_{j=1}^{n} \lambda_j x_{ij} = \phi x_{io} \quad i = 1,\ldots, m; \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} \quad r = 1,\ldots, s; \qquad \text{(VW)} \\
&\sum_{j=1}^{n} \lambda_j = 1; \\
&\lambda_j \ge 0 \ \ and \ \ \phi \ \ free.
\end{aligned}$$

The only difference between (VS) and (VW) is that the input inequalities are changed to equalities. The referent frontier DMUs are those in the basis when calculating the strong disposability model, say (VS). If we have some referent DMUs such that no non-zero input slack values are detected for DMU_o, then we have, at optimality,

$$\sum_{j \in B} \lambda_j^* x_{ij} = \theta^* x_{io}$$

where B represents the set of referent DMUs, $B=\{j|\ \lambda_j^* > 0\}$. Obviously, λ_j^* and θ^* are also optimal for VW, therefore $\theta^* = \phi^*$. Thus no input congestion occurs. This completes the *if* part.

To establish the *only if* part, we note that if no input congestion occurs, then at optimality, there exists a basis B' such that $\sum_{j \in B'} \lambda_j^* x_{ij} = \phi^* x_{io} = \theta^* x_{io}$. This indicates that there exist some referent DMUs such that the input constraints are binding in (VS). Therefore no non-zero input slack values are detected by reference to those DMUs in B'. –

[Remarks]: The input slacks in (VS) do not necessarily represent DEA slack values (Banker *et al.* 1984). However, if all frontier DMUs are extreme efficient, e.g., A and B, in Figure 1.A, then the input slacks are the same as the DEA slack values. In the right Figure 1, because C can be compared to a convex combination of D (weakly efficient) and A, no input slack is announced. (With a DEA model, the same DEA slack should be obtained in either the left or right Figure 1.) See Cooper, Seiford and Zhu (2000) for discussions on congestion measures based upon slacks.

It is well know that in the single input and the single output situation, no input or output slack will occur for CRS measures, whereas, non-zero slack values may occur for VRS measures. That is to say, congestion will never occur with CRS but can possibly happen with VRS. Clearly for the single input - single output case, the congestion component depends upon the RTS assumption. Thus, in general one should not be surprised to obtain different congestion results from different RTS assumptions.

In the authors' experience that the DEA efficient frontiers of most real world data sets are composed solely of extreme efficient DMUs. Therefore, we readily have the following:

Corollary: If the efficient frontier is only composed of extreme efficient DMUs, then congestion occurs *if and only if* non-zero slack values are detected. Furthermore, the factors responsible for the congestion are those with non-zero slack values.

Färe *et al.* (1994) introduced a procedure for detecting the factors responsible for the congestion. By the above results (non-zero slack values), one can easily find and identify congestion and its sources without the need for calculating another corresponding DEA measure satisfying weak disposability.

Finally to address McDonald's (1996) concern, as to the order of computation of the scale and congestion components, we have that

Decomposition 1 (definition 1): $\dfrac{CS}{VW} = \dfrac{CS}{VS}\dfrac{VS}{VW} = \dfrac{VS}{VW}\dfrac{CS}{VS}$

Decomposition 2 (definition 2): $\dfrac{CS}{VW} = \dfrac{CW}{VW}\dfrac{CS}{CW} = \dfrac{CS}{CW}\dfrac{CW}{VW}$

It is obvious from the above two equations that the outcomes have nothing to do with the computation orders but rather reflect the different technologies upon which the computations are based.

3. Conclusion

We have examined the decomposition of inefficiency into scale and congestion components and have shown that those two components depend upon both disposability and RTS assumptions. Caution should be exercised when discussing a particular decomposition. McDonald (1996) proposes two different decompositions and obtains dramatically opposite outcome. He erroneously concludes the decomposition is arbitrary and misleading. In fact, the different outcomes of scale and congestion measures should be expected because different technologies were assumed.

In closing, we note that the results and framework of his paper can be further extended in several obvious directions. Scale component could also be defined under different output disposability assumptions. Congestion measures could be defined under other RTS assumptions such as nonincreasing or nondecreasing returns to scale. Again different should be expected from those new technologies.

Finally, we limited our discussion of decomposition to the input-oriented DEA measures. Similar remarks can be made for the output-oriented DEA measures.

References

Banker, R.D. and R.M. Thrall, "Estimation of Returns to Scale using Data Envelopment Analysis," *European J. of Operational Research*, 62(1992), 74-84.

Banker, R.D., W.W. Cooper and A. Charnes, "Some Models for Estimating Technical and Scale Inefficiencies in DEA," *Management Sci.*, 30 (1984), 1078-1092

Byrnes, P., R. Färe and S. Grosskopf, "Measuring Productive Efficiency: Application to Illinois Strip Mines," *Management Sci.*, 30 (1984), 671-681.

Cooper, W.W., Seiford, L.M. and Zhu, Joe, A unified additive model approach for evaluating inefficiency and congestion with associated measures in DEA. *Socio-Economic Planning Sciences*, Vol. 34, No. 1 (2000), 1-25.

Färe, R., S. Grosskopf and C.A.K. Lovell, *Production Frontiers*, Cambridge University Press, Cambridge, 1994.

McDonald, J., "A Problem with the Decomposition of Technical Inefficiency into Scale and Congestion Components," *Management Sci.*, 42(1996), 473-474.

Seiford, L.M. and Zhu, Joe, An investigation of returns to scale under data envelopment analysis. *OMEGA, International Journal of Management Science*, Vol. 27, No. 1 (1999), 1-11.

Zhu, Joe and Z. Shen, "A Discussion of Testing DMUs' Returns to Scale," *European J. of Operational Research*, 81(1995), 590-596.

An Approach for Determining DEA Efficiency Bounds

Yao Chen, Hiroshi Morita, Joe Zhu[1]

College of Management, University of Massachusetts at Lowell, Lowell, MA 01845 USA
Graduate School of Information Science and Technology, Osaka University, Suita, Osaka, 565-0871, Japan
Department of Management, Worcester Polytechnic Institute, Worcester, MA 01609 USA

Abstract: Constrained facet analysis is used to evaluate decision making units (DMUs) which have non-zero slacks in data envelopment analysis (DEA) by requiring a full dimensional efficient facet (FDEF). The current paper shows that the FDEF-based approach may deem those extreme efficient DMUs which are not located on any FDEF as inefficient. Using strong complementary slackness condition (SCSC) solutions, this paper develops an alternative method for the treatment of non-zero slack values in DEA. The newly proposed method can deal with the situation when FDEFs do not exist.

Keywords: Data Envelopment Analysis (DEA); Efficient; Slack; Strong Complementary Slackness Condition (SCSC).

1. Introduction

Experience with the application of data envelopment analysis (DEA) shows that two inefficient decision making units (DMUs) may have the same efficiency score, but one may have larger amount of underutilized resources or unachieved outputs, i.e., non-zero slacks, than the other. Those non-zero slacks are treated by an infinitesimal (ε) Charnes et al. (1979) (CCR), or by extrapolated efficient facets in Bessent et al. (1988) and Chang and Guh (1991).

The constrained facet analysis by Bessent et al. (1988) fails to work when extreme efficient DMUs span a non full dimensional efficient facet (FDEF). A FDEF refers to an efficient facet with $m+s-1$ dimension under CCR model, where m and s are the numbers of inputs and outputs, respectively. Green et al. (1996)

[1] Joe Zhu wants to thank the financial support from the Japan Society for Promotion of Science (JSPS). The paper was finished while Joe Zhu was visiting the Osaka University under the JSPS Invitation Research Fellowship.

develop a mixed-binary linear programming problem to treat the non-zero slacks. Green et al. (1996) claim that a basic requirement for their implementation is that there must be at least $m+s-1$ CCR extreme efficient DMUs. However, $m+s-1$ CCR extreme efficient DMUs do not necessarily span a FDEF, and they may span several non-FDEFs. Particularly, some extreme efficient DMUs which are not located on any FDEF may be termed as inefficient by the mixed-binary linear programming problem even the number of extreme efficient DMUs is greater than $m+s-1$. This indicates that Green et al.'s (1996) new implementation also requires the existence of FDEFs, rather than the existence of at least $m+s-1$ extreme efficient DMUs. But this condition may not be satisfied in real world applications (see Bessent et al. 1988).

The current paper places lower bounds that are obtained by strong complementary slackness condition (SCSC) solution pairs for extreme efficient DMUs. It is shown that our method does not change the efficiency ratings for those DMUs which do not have non-zero slacks.

2. Determination of the lower bounds

Suppose we have n decision making units (DMUs). Each DMU_j $j = 1, 2, \ldots, n$ produces s different outputs y_{rj} $(r = 1, 2, \ldots, s)$ using m different inputs x_{ij} $(i = 1, 2, \ldots, m)$. Then the Charnes et al. (1978, 1979) (CCR) model with infinitesimal ε can be expressed as:

$$\begin{aligned} &\max \sum_{r=1}^{s} u_r y_{ro} \\ &s.t. \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{s} v_i x_{ij} \le 0 \qquad j = 1, \ldots, n; \\ &\qquad \sum_{i=1}^{m} v_i x_{io} = 1 \\ &\qquad u_r, v_i \ge \varepsilon. \end{aligned} \tag{1}$$

The above CCR model is equivalent to the following fractional programming model:

$$\begin{aligned} &\max h_o = \frac{\sum_{r=1}^{s} \mu_r y_{ro}}{\sum_{i=1}^{m} w_i x_{io}} \\ &s.t. \quad h_j = \frac{\sum_{r=1}^{s} \mu_r y_{rj}}{\sum_{i=1}^{m} w_i x_{ij}} \le 1 \quad j=1,2,\ldots,n \\ &\qquad \frac{\mu_r}{\sum_{i=1}^{m} w_i x_{io}} \ge \varepsilon, \frac{w_i}{\sum_{i=1}^{m} w_i x_{io}} \ge \varepsilon. \end{aligned} \tag{2}$$

On the basis of CCR model (1), all n DMUs can be partitioned into six classes E, E′, F, NE, NE′, and NF (Charnes et al. 1991). The first two are efficient, F class is frontier but has non-zero slacks, and the last three are DEA projections in E, E′, F, respectively.

DMUs in class E determine the efficient facets. An efficient facet can be either a FDEF spanned by $m+s$-1 CCR extreme efficient DMUs, or a non-FDEF spanned by some CCR extreme efficient DMUs whose number is less than $m+s$-1. Here we need a regularity condition that every subset of $m+s$-1 extreme efficient DMUs is linearly independent. Otherwise, if a subset of $m+s$-1 extreme efficient DMUs is linearly dependent, then the dimension of this efficient face will be strictly less than $s+m$-1.

Note that the purpose of introducing non-Archimedean ε is to impose the positivity on DEA multipliers. The non-zero slacks are relative to the zero optimal values of multipliers. If we can determine the positive lower bound on each multiplier, then by the complementary slackness condition of linear programming, we can suppress the non-zero slacks and consequently obtain a comparable overall efficiency score for each DMU.

Note that there must exist a non-zero optimal multiplier solution for a DMU in set E. But we cannot conclude that for a DMU in set E, the optimal multipliers are always positive after running the DEA model without ε. Because of the multiple optima, zero optimal multipliers are likely to occur. In order to deal with this situation, we use the solutions that satisfy strong complementary slackness condition (SCSC) which states that there exists an optimal solution (λ_j^*, s_r^+, s_i^- u_r^*, v_i^*, t_j^*) for which, in addition to complementary slackness condition, we have $s_r^+ + u_r^* > 0$ (r = 1, ..., s), $s_i^- + v_i^* > 0$ (i = 1, ..., m), $\lambda_j^* + t_j^* > 0$ (j = 1, ..., n), where $t_j^* = -\sum_{r=1}^{s} u_r y_{rj} + \sum_{i=1}^{s} v_i x_{ij}$.

Lemma 1: For a specific DMU_o, let s_i^-, s_r^+, u_r^o and v_i^o be an optimal solution. If this optimal solution satisfies SCSC, then u_r^o and v_i^o are all positive.

On the basis of Lemma 1, we can find a set of positive optimal dual multipliers for each DMU in set E. As stated in Charnes et al. (1991), the procedure for computing SCSC solution is well adapted for DMUs in set E. But first we need to find out DMUs in set E. By the resent results of Thrall (1996), we solve the following modified DEA model (see Seiford and Zhu (1999) for a detailed discussion on this type of DEA models):

$$\begin{aligned} &\max \sum_{r=1}^{s} u_r y_{ro} \\ &s.t. \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{s} v_i x_{ij} \le 0 \qquad j \ne o; \\ &\qquad \sum_{i=1}^{m} v_i x_{io} = 1 \\ &\qquad u_r, v_i \ge 0. \end{aligned} \tag{3}$$

The DMU_o under evaluation is excluded from reference set. Then $DMU_o \in$ E if and only if the optimal value to (3) is greater than one, or (3) is infeasible.

Next, since we only interest in positive multipliers, we find a SCSC solution for a specific $DMU_o \in$ E by the following model:

$$\begin{aligned} &\max \sum_{j\neq o} t_j + \sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i \\ &s.t. \quad -\sum_{r=1}^{s} u_r y_{rj} + \sum_{i=1}^{s} v_i x_{ij} = t_j \qquad j \neq o; \\ &\sum_{i=1}^{m} v_i x_{io} = 1 \\ &\sum_{r=1}^{s} u_r y_{ro} = 1 \\ &u_r, v_i, t_j \geq 0. \end{aligned} \tag{4}$$

We solve (4) and remove positive t_j, u_r, and v_i, then re-run (4) until the optimal value is zero, i.e., all t_j, u_r, and v_i are removed from the objective function. An optimal multiplier solution pair (u_r^o, v_i^o) that satisfies SCSC is the average of all (u_r, v_i) in each step.

Now for convenience, let first Q DMUs be those in set E, i.e., $DMU_q \in$ E for q =1, 2, ..., Q. Denote SCSC solution for each DMU_q as u_r^q and v_i^q

Then we have the following algorithm:

Step 1: Find out all extreme efficient DMUs by (3).

Step 2: Compute SCSC solution for each DMU in set E by (4).

Step 3: Let $u_r^* = \min_{q=1,\dots,Q} \{u_r^q\}$ and $v_i^* = \min_{q=1,\dots,Q} \{v_i^q\}$.

Step 4: Denote set $I = \{i \mid x_{ij} \text{ has zero slack value}, DMU_j \in F \text{ or } NF\}$ and set $R = \{r \mid y_{rj} \text{ has zero slack value}, DMU_j \in F \text{ or } NF\}$.

Step 5 Add the constraints $v_i \geq v_i^*$ $i \in I$ and $u_r \geq u_r^*$ $r \in R$ into the unbounded ($\varepsilon = 0$) DEA fractional programming model:

$$\begin{aligned} &\max \frac{\sum_{r=1}^{s} \mu_r y_{ro}}{\sum_{i=1}^{m} w_i x_{io}} \\ &s.t. \ \frac{\sum_{r=1}^{s} \mu_r y_{rj}}{\sum_{i=1}^{m} w_i x_{ij}} \leq 1 \quad \text{j} = 1,2,\dots, \text{n} \\ &\mu_r \geq u_r^*, r \in R; w_i \geq v_i^*, i \in I. \end{aligned} \tag{5}$$

Step 6: Run the DEA model (5) with lower weight bounds.

Note that the lower bound DEA model (5), is a fractional programming problem. By the Charnes-Cooper transformation in Charnes and Cooper (1962), we can obtain the following equivalent linear multiplier DEA model.

$$\max \sum_{r=1}^{s} u_r y_{ro}$$
$$\begin{aligned} s.t.\quad & \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{s} v_i x_{ij} \le 0 \qquad j = 1,\dots,n; \\ & \sum_{i=1}^{m} v_i x_{io} = 1 \\ & -u_r + t_o u_r^* \le 0 \qquad r \in R \\ & -v_i + t_o v_i^* \le 0 \qquad i \in I \\ & u_r \ge 0, r \notin R; v_i \ge 0, i \notin I. \end{aligned} \tag{6}$$

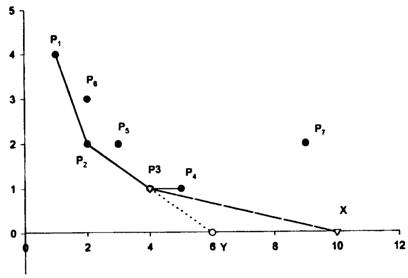
in which $t_o = (\sum_{i=1}^{m} w_i x_{io})^{-1} > 0$ is the transformation factor for DMU_o (see also Roll et al. 1991). It is easy to show the following.

Lemma 2: Suppose u_r^o and v_i^o is an optimal solution to (1) with $\varepsilon = 0$, then it is also an optimal solution to (2) with $\varepsilon = 0$.

Theorem: (i) If $DMU_o \in$ E, or E', or NE, or NE' under the DEA model without ε, then the efficiency classification of this DMU_o remains the same under the DEA model with lower weight bounds of u_r^* and v_i^*, (ii) If $DMU_o \in$ F or NF under the DEA model without ε, then the efficiency score of DMU_o will be changed under the DEA model with lower weight bounds of u_r^* and v_i^*.

The above theorem indicates that (i) the approach described here not only allows the weight flexibility in original DEA model, but also imposes positivity of multipliers so that we obtain an efficient frontier rather than a weakly efficient frontier, and (ii) our approach does not require the existence of FDEF.

The proposed technique is demonstrated by seven DMUs given in Chang and Guh (1991) (see Figure 1).



	P_1	P_2	P_3	P_4	P_5	P_6	P_7
output 1, y:	1	1	1	1	1	1	1
input 1, x_1:	1	2	4	5	3	2	9
input 2, x_2:	4	2	1	1	2	3	2

Fig. 1. Hypothetical Efficient Facet

We have that E={P_1, P_2, P_3}, NE'={P_5, P_6}, F={P_4} and NF={P_7}. The SCSC solutions for P_1, P_2, and P_3 are (1, 0.60, 0.10), (1, 0.25, 0.25), and (1, 0.10, 0.60), respectively

Thus $u^* = 1, v_1^* = 0.1, v_2^* = 0.1$ and I = {1} and R = ∅. Then we have the lower bounds of $w_1 \geq 0.1$ in model (7). Now we obtain the adjusted efficiency scores 10/11 and 10/21 respectively for P_4, and P_7.

However, we should note that our approach is dependent of the choice of SCSC solutions. As a matter of fact, the lower bounds (SCSC solutions) introduce some hypothetical efficient facets at the end of original efficient facets. These kinds of facets are part of different supporting hyperplanes at DMUs in set E. Thus, if we choose a different set of SCSC solutions, the lower bounds will be changed and consequently, the adjusted efficiency score for a DMU∈F or NF may be different. For example, the lower bound in the above example introduce a new point X (10,0) and P_3X becomes our hypothetical extended efficient facet for measuring the efficiency of P_4 and P_7 which have non-zero slack on x_1 (see Figure 1). If there exists a FDEF near the area of P_4 and P_5, say P_2 P_3, then P_3Y is the extrapolated efficient facet by Chang and Guh (1991) and Bessent et al. (1988). Different SCSC solutions may establish different hypothetical efficient facets in the region below ray P_3 P_4 and above P_3Y. In fact, the infinitesimal ε constructs a hypothetical efficient facet with norm (ε, 1) in that region for P_4 and P_7.

References

Bessent, A., Bessent, W., Elam, J., and Clark, T. (1988), "Efficiency frontier determination by constrained facet analysis", *Operations Research* 36/5, 785-796.

Chang, K.P., and Guh, Y.Y. (1991), "Linear production functions and data envelopment analysis ", *European Journal of Operational Research* 52, 215-223.

Charnes, A., and Cooper, W.W. (1962), "Programming with linear fractional functions", *Naval Res. Logist. Quart.* 9, 181-186.

Charnes, A., Cooper, W.W., and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research* 2/6, 429-444.

Charnes, A., Cooper, W.W., and Thrall, R.M. (1991), "A structure for classifying and characterizing efficiency and inefficiency in data envelopment analysis", *Journal of Productivity Analysis* 2, 197-237.

Green, R.H., Doyle, J.R., and Cook, W.D. (1996), "Efficiency bounds in data envelopment analysis", *European Journal of Operational Research* 89, 482-490.

Roll, Y., Cook, W.D. and Golany, B. (1991), "Controlling factor weights in data envelopment analysis", *IIE Transaction* 23, 2-9.

Seiford, L.M. and J. Zhu (1999), "Infeasibility of super-efficiency data envelopment analysis models," *INFOR* **37** (May) 174-187.

Thrall, R. M. (1996), "Duality, classification and slacks in DEA," *Annals of Operations Research* **66** 109-138.

An Extended Approach of Multicriteria Optimization for MODM Problems

Hua-Kai Chiou [a], Gwo-Hshiung Tzeng [b,1]

[a] Institute of Management of Technology, National Chiao Tung University; Faculty of Department of Statistics, National Defense University, P.O. Box 90046-15, Chungho, 235, Taiwan. Fax:+886-2-22250488

[b] Institute of Management of Technology, National Chiao Tung University, 1001, Ta-Hsuch Rd., Hsinchu 300, Taiwan. Fax: +886-3-5726749

Abstract

In this paper we propose an extended method for multicriteria optimization and compromise solution to solve multiple objective decision making (MODM) problems. This method assumes that optimal compromise solution should have the shortest distance from the positive ideal solution (PIS) as well as the longest distance from the negative ideal solution (NIS). We use the membership function of fuzzy set theory to express the satisfaction level, and use max-min operation for this bi-objective programming problem. To illustrate this procedure, prequalification for the project bidding process of an outsourcing partner for semiconductor enterprise in Taiwan is solved by use of our procedure.

Keywords: compromise solution, fuzzy set theory, max-min operation, semiconductor, prequalification

1. Introduction

Kuhn and Tucker (1951) published one of earliest considerations of multiple objectives using vector optimization concept, followed by Yu (1973) who proposed a compromise solution method for coping with multicriteria decision-making (MCDM) problems. Subsequently, there have many works using MCDM for applications such as transportation investment and planning, econometric and development planning, capital budgeting, investment portfolio selecting, health care planning, forest management, public policy and environmental issues, etc.

Dealing with MCDM problems decision makers have more than one objective or goal in selecting a course of action, while satisfying the constraints dictated by

[1] Corresponding Author, National Distinguished Chair Professor.
E-mail: ghtzeng@cc.nctu.edu.tw (G H. Tzeng); hkchiou@rs590.ndmc.edu.tw (H K. Chiou)

environment, process and resources. Mathematically, these problems can be represented as:

$$\begin{aligned} &\max/\min \ [f_1(x), f_2(x),\ldots, f_k(x)] \\ &\quad \text{s.t.} \quad x \in X, X = \{x \mid g_i(x) \le 0,\ i = 1,2,\ldots,m\} \end{aligned} \tag{1}$$

where x is an n-dimensional vector of decision variables, that consists of m constraints and k objectives. Any or all of the functions may be nonlinear.

On account of incommensurability and conflicting nature of the multiple criteria, Yu (1973) proposed compromise solution methods. Hwang and Yoon (1981) proposed TOPSIS using the concept of optimal compromise solution. Lai et al. (1994) further utilized the Euclidean distance to drive the TOPSIS approach for MODM problem. In addition, Opricovic (1998) proposed a new approach called VIKOR to solve the MODM problem. However, the VIKOR method considered only the shortest distance of L_p-metric from PIS.

In this paper we extend both VIKOR and the compromise solution method, also considering the optimal compromise solution that satisfies both the shortest distance from PIS and the longest distance from NIS. We also use membership functions to express the satisfaction level and employ max-min operation for this bi-objective programming problem. To illustrate this procedure, the selection of an outsourcing partner for a Taiwan semiconductor enterprise is solved by use of this procedure.

Follows, the concept of multicriteria ranking and compromise solution with distance of Minkowski's L_p-metric is reviewed in Section 2; the evaluation model of this study is presented in Section 3; an illustrative example shown in Section 4 demonstrates this model in practice; the conclusions are summarized in Section 5.

2. The Multicriteria Metric for Compromise Ranking Methods

With a given reference point, the MODM problem can then be solved by locating decision points that are the closest to the reference point. Generally, the global criteria method measures the distance using Minkowski's L_p-metric, which defines the distance between two points, f and f^* in k-dimensional space as follows

$$L_p = \left\{ \sum_{i=1}^{k} (f_i^* - f_i)^p \right\}^{1/p}, \text{ where } p \ge 1 \tag{2}$$

where f^* is the reference point, and distance L_p decreases as p increases, i.e., $L_1 \ge L_2 \ge \cdots \ge L_\infty$. Specifically, L_1, called the Manhattan distance and L_2, called the Euclidean distance are the longest and the shortest distances in the geometrical sense, respectively, whereas L_∞, called the Tchebycheff distance, is the shortest distance in the numerical sense, that is,

$$L_\infty = \max_i \left\{ \left| f_i^* - f_i \right| \right\} \tag{3}$$

With the concept of compromise solution, we then transfer Eq.(1) to the following bi-objective problem,

$$\begin{aligned} \min \quad & L_p^{PIS}(x) \\ \max \quad & L_p^{NIS}(x) \\ \text{s.t.} \quad & x \in X \end{aligned} \tag{4}$$

where $p = 1,2,\ldots,\infty$, L_p^{PIS} and L_p^{NIS} represent the distance of L_p-metric from PIS and from NIS, respectively. Since there are usually conflicts with each objective, it is difficult to simultaneously obtain their individual optima. Thus, we utilize the membership functions $\mu_1(x)$ and $\mu_2(x)$ to represent the satisfactory level of bi-objective functions and use the max-min operation (Bellman and Zadeh 1970, Zimmermann 1978) to drive the equivalent model giving the same values of α:

$$\begin{aligned} \max \quad & \alpha \\ \text{s.t.} \quad & \mu_1(x) \geq \alpha \ \ and \ \ \mu_2(x) \geq \alpha \\ & x \in X \end{aligned} \tag{5}$$

where $\alpha = \min(\mu_1, \mu_2)$ is the minimal satisfactory level for both objectives, and in practice the parameter α is generally subjectively selected by DM.

3. The Extended Compromise Ranking Approach

In this study, we consider that the optimal compromise solution should have the shortest distance from PIS as well as the longest distance from NIS. Firstly, we employ normalization by reference point to remove the effects of the incommensurability nature, and then establish the algorithm as following steps:

1. Determine the best value f_i^* and the worst value f_i^- of all criterion functions, where f_{ij} is the value of *i*-th criterion function for the *j*-th alternative, respectively, for criteria *i*=1,…,*k*, we have

   $f_i^* = \max_j f_{ij}$, $\quad f_i^- = \min_j f_{ij}$

2. Compute the values of S_j^{PIS}, S_j^{NIS}, R_j^{PIS}, and R_j^{NIS} for $j = 1,\ldots,J$. The first two represent the L_p-metric for $p = 1$ and the last two represent the L_p-metric for $p = \infty$ from PIS and from NIS, respectively.

$$\begin{aligned} S_j^{PIS} &= \sum_{i=1}^{k} w_i \cdot \frac{f_i^* - f_{ij}}{f_i^* - f_i^-}; & S_j^{NIS} &= \sum_{i=1}^{k} w_i \cdot \frac{f_{ij} - f_i^-}{f_i^* - f_i^-}; \\ R_j^{PIS} &= \max_i \left[w_i \cdot \frac{f_i^* - f_{ij}}{f_i^* - f_i^-} \right]; & R_j^{NIS} &= \max_i \left[w_i \cdot \frac{f_{ij} - f_i^-}{f_i^* - f_i^-} \right] \end{aligned} \tag{6}$$

where w_i represents the relative weight of i-th criterion.

3. Compute the values Q_j^{PIS} and Q_j^{NIS} for $j=1,\dots,J$ which is defined as

$$Q_j^{PIS} = v_P \cdot \frac{S_j^{PIS} - (S^{PIS})^-}{(S^{PIS})^* - (S^{PIS})^-} + (1-v_P)\cdot \frac{R_j^{PIS} - (R^{PIS})^-}{(R^{PIS})^* - (R^{PIS})^-}$$
$$Q_j^{NIS} = v_N \cdot \frac{S_j^{NIS} - (S^{NIS})^-}{(S^{NIS})^* - (S^{NIS})^-} + (1-v_N)\cdot \frac{R_j^{NIS} - (R^{NIS})^-}{(R^{NIS})^* - (R^{NIS})^-} \tag{7}$$

where $0 \le v_p \le 1$, $0 \le v_N \le 1$, and

$(S^{PIS})^* = \max_j S_j^{PIS}; (S^{PIS})^- = \min_j S_j^{PIS}; (R^{PIS})^* = \max_j R_j^{PIS}; (R^{PIS})^- = \min_j R_j^{PIS};$

$(S^{NIS})^* = \max_j S_j^{NIS}; (S^{NIS})^- = \min_j S_j^{NIS}; (R^{NIS})^* = \max_j R_j^{NIS}; (R^{NIS})^- = \min_j R_j^{NIS}.$

4. The objective of our approach is to solve following mathematic programming:

$$\begin{aligned} \min \quad & Q_j^{PIS}(x) - Q_j^{NIS}(x) \\ \text{s.t.} \quad & x \in X \end{aligned} \tag{8}$$

where X is the set of feasible solutions, and setting the same importance for the values Q_j^{PIS} and Q_j^{NIS} in this study.

4. Illustrative Example

Project mangers are faced with decision environments in complex projects. The elements of the problems are numerous; the interrelationships among the elements are extremely complicated; and human value and judgment systems are integral elements of project problems (Lifson and Shaifer, 1982). Therefore, the ability to make sound decisions is very important to the success of a project.

Prequalification is defined by Stephen (1984), Moore (1985) and Clough (1986) as the screening of construction contractors by project owners or their representatives according to a predetermined set of criteria deemed necessary for successful project performance. Thus, prequalification means that the contracting firm wishing to bid on a project needs to be qualified before it can be issued bidding documents on which it can submit a proposal. Prequalification can aid public and private owners in achieving successful and efficient use of their funds by ensuring that only qualified contractors will bid on the project.

A simplified project example of contractor prequalification is presented here for demonstration purposes. To simplify calculations, the seven factors that are considered to evaluate this project example for prequalification are experience, financial stability, capital assets, quality performance, manpower resources, equipment resources and current workload.

Table 1 presents a project example for which contractors A, B, C, D and E wish to prequalify. In order to evaluate these participating contractors, we first employ AHP to aggregate the judgment of group decision-making behavior. After deriving the relative weight among considered criteria shown as Table 2.

Furthermore, we utilize the extended compromise ranking method step by step in Section 3, where there are two minimal criteria, financial stability and current workload. For both of these, smaller values are better, while other criteria seek maximal values; here we pick $v_P = v_N = 0.5$ in general. Finally we can conduct the Q_j^{PIS} and Q_j^{NIS} by Eq.(7), and obtain the preference order of participating contractors by Eq.(8), as shown in Table 3: $E \succ C \succ D \succ B \succ A$, where $E \succ C$ indicates the prequalification result of contractor *E* being superior to contractor *C*. This has almost the same preferential order with VIKOR method, except for contractor *C* and contractor *D*. It seems that these two contractors are not comparable because their difference in *Q* value (Opricovic 1998) does not exceed 0.25.

Table 1. Qualification of participating contractors in project bid

Contractor	A	B	C	D	E
Experience	5 years	7 years	8 years	10 years	15 years
Financial Stability	High growth rate no liability	$5.5 M liabilities	$6 M liabilities	$4 M liabilities	$1.5 M liabilities
Capital Asset	$7 M assets	$10 M assets	$14 M assets	$11 M assets	$6 M assets
Quality Performance	Good	Medium	Good	Good	Weak
Manpower Resources	150 laborers	100 laborers	120 laborers	90 laborers	40 laborers
Equipment Resources	4 mixer machines	6 mixer machines	1 mixer machines	4 mixer machines	2 mixer machines
Current Workload	2 project in mid and 1 big project ending	2 projects ending	1 medium project started and 2 projects ending	1 medium project in mid and 2 big projects ending	2 small projects started and 3 projects ending

Table 2. The weight of considered criteria by AHP

Factor	Experience	Financial Stability	Capital Assets	Quality Performance	Manpower Resources	Equipment Resources	Current Workload
weight	0.372	0.204	0.102	0.148	0.053	0.039	0.082

Table 3. The compromise solutions' distance from PIS and NIS w.r.t. each contractor

Contractor	A	B	C	D	E
Value of Q_j^{PIS}	1.00000	0.73858	0.31201	0.27693	0.00000
Value of Q_j^{NIS}	0.07088	0.09535	0.50976	0.45157	1.00000
$Q_j^{PIS} - Q_j^{NIS}$	0.92912	0.64323	-0.19775	-0.17465	-1.00000
Preferential order	*5*	*4*	*2*	*3*	*1*

5. Conclusion

Project management involves complex decision-making situations that require discerning abilities and methods to make sound decisions. This paper has successfully demonstrated the revised compromise ranking method using L_p-metric distance family to find optimal solutions that have shortest distance from positive ideal solution, as well as has longest distance from negative ideal solutions.

In addition, this compromise solution is stable within different decision-making processes, whether they be "voting by majority rule" when either $v_P > 0.5$ or $v_N > 0.5$ is needed, or "by consensus" with either $v_P \approx 0.5$ or $v_N \approx 0.5$, or "with veto" with either $v_P < 0.5$ or $v_N < 0.5$. Both v_P and v_N are the weights of decision making strategy with the majority of criteria.

References

[1]Bellman, R.E., and Zadeh, L.A.(1970) "Decision-making in a fuzzy environment", *Management Science*, **B17**(1), 141-164.

[2]Clough, R.(1986) *Construction Contracting*, New York, Wiley.

[3]Hwang, C.L., and Yoon, K.(1981) *Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, Heidelberg.

[4]Kuhn, H.W., and Tucker, A.W.(1951) "Nonlinear programming", *Proceedings 2nd Berkeley Symposium on Mathematical Statistics and Probability*, 481-491. J. Neyman (Ed.), University of California Press, Berkeley.

[5]Lai, Y.J., Liu, T.Y., and Hwang, C.L.(1994) "TOPSIS for MODM", *European Journal of Operational Research*, **76**(3), 486-500.

[6]Lifson, M.W. and Shaifer, E.F.(1982), *Decision and Risk Analysis for Construction Management*, New York, Wiley.

[7]Moore, M.J.(1985) "Selecting a contractor for fast-track projects: Part I, principles of contractor evaluation", *Plant Engineering*, **39**(1), 74-75.

[8]Opricovic, S.(1998) *Multicriteria Optimization in Civil Engineering*, Faculty of Civil Engineering, Belgrade.

[9]Stephen, A.(1984) *Contract Management Handbook for Commercial Construction*, CA: Naris Publications.

[10]Yu, P.L.(1973) "A class of solutions for group decision problems", *Management Science*, **19**(8), 936-946.

[11]Yu, P.L.(1985) *Multiple Criteria Decision Making: Concepts, Techniques, and Extensions,* Plenum, New York.

[12]Zimmermann, H.J.(1978) "Fuzzy programming and linear programming with several objective functions", *Fuzzy Sets and Systems*, **1**(1), 45-55.

The Method of Elastic Constraints for Multiobjective Combinatorial Optimization and its Application in Airline Crew Scheduling

Matthias Ehrgott and David M. Ryan

Department of Engineering Science, University of Auckland, Private Bag 92019, Auckland, New Zealand, email {m.ehrgott, d.ryan}@auckland.ac.nz

Abstract. We consider scalarization techniques for multiobjective combinatorial optimization. We briefly discuss the problems occurring in known methods and show that a new method of elastic constraints can overcome these. The method is a generalization of both the weighted sum method and the ε-constraint method and generates all Pareto optimal solutions. We show an application of the method in airline crew scheduling, where we optimize cost and robustness of solutions. Numerical results on real world instances are given.

1 Multiobjective Combinatorial Optimization

A multiobjective combinatorial optimization problem (MOCO) is the following mathematical program:

$$\begin{aligned} \min(c_1^T x, \ldots, c_Q^T x) & \\ \text{subject to} \quad Ax &\le b \\ x &\ge 0, \text{ integer.} \end{aligned} \tag{1}$$

Here $c_j \in \mathbb{R}^n$, $j = 1, \ldots, Q$ are Q objective vectors, $x \in \mathbb{R}^n$ is a vector of variables, $A \in \mathbb{R}^{m \times n}$ is a constraint matrix and $b \in \mathbb{R}^m$ a right hand side vector. It is therefore an integer programming problem with n variables x_i (these are usually binary), m constraints, and Q linear objective functions. The feasible set $X = \{x : Ax \le b; x \ge 0; x \text{ integer}\}$ is finite and represents a combinatorial structure, e.g. the set of perfect matchings of a graph. We understand the minimization in the sense of Pareto optimality (efficiency). For a survey on the state of the art in multiobjective combinatorial optimization see [2].

MOCO problems are usually solved using scalarization techniques. The most popular one of weighted sum scalarization is

$$\min_{x \in X} \sum_{i=1}^{Q} \lambda_i c_i^T x, \tag{2}$$

which uses a weighting vector $\lambda \in \mathbb{R}_+^Q$ with $\sum_{i=1}^Q \lambda_i = 1$. It has the drawback that it cannot find all Pareto optimal solutions, due to the feasible set X not being convex. On the other hand, it preserves the structure of the problem and the computational effort for solving the scalarized problem is exactly the same as for the single objective version of the problem under consideration.

The situation is completely different for the ε-constraint method, which is based on the scalarization

$$\begin{aligned} &\min_{x \in X} c_i^T x \\ &\text{subject to } c_j^T x \leq \varepsilon_j, \quad j \neq i, \end{aligned} \tag{3}$$

where the ε_j denote bounds on the objectives $c_j^T x$. Using this, all Pareto optimal solutions can be generated. However, the knapsack constraints that have been added imply that the scalarized problem is often NP-hard, even if the single objective version of the problem is polynomially solvable, see references in [2]. In addition, the constraints also tend to make NP-hard problems even harder, as we shall see below.

Similar comments apply for other methods, such as the augmented weighted Chebychev method, or Benson's method. The reason is that the scalarizations used in these methods are essentially models in which constraints on the objectives are present.

2 The Method of Elastic Constraints

From the above observations the challenge is then to find a solution method for MOCO problems in which the scalarizations do not introduce too much additional difficulty, like the ones based on constraints, and which are able to generate the whole set of Pareto optimal solutions.

Such a technique is the method of elastic constraints. The idea is to allow the constraints on objective values in (3) to be violated, but to penalize that violation in the objective function. Thus the scalarization becomes

$$\begin{aligned} &\min_{x \in X} c_i^T x + \sum_{j \neq i} p_j su_j \\ &\text{subject to } c_j^T x + sl_j - su_j = \varepsilon_j \quad j \neq i. \\ &\qquad\qquad sl_j, su_j \geq 0 \quad j \neq i \end{aligned} \tag{4}$$

Here the constraints of (3) are relaxed by introducing a slack variable sl_j and a surplus variable su_j for each of the constraints on objective values. Parameters p_j penalize positive values of su_j, i.e. constraint violation. The following theorem is the main result about this method.

Theorem 1. *1. An optimal solution of (4) is a Pareto optimal solution of (1), if $p_j > 0$ for all $j \neq i$.*
2. Let x^* be a Pareto optimal solution of (1). Then there is some $\varepsilon \in \mathbb{R}^Q$ and for each $i = 1, \ldots, Q$ a vector $p^i \in \mathbb{R}^{Q-1}$ such that x^* defines an

optimal solution of (4) for this i for all penalty vectors $p \in I\!R^Q$ with $p_j \geq p_j^i, j \neq i$.

Note that the first part is trivial. For the second it is sufficient to consider $\varepsilon = (c_1^T x^*, \dots, c_Q^T x^*)$ and $p_j^i = \max(\max_{\{l: c_i^T x^l \leq c_i^T x^*\}} (c_i^T x^* - c_i^T x^l)/(c_j^T x^* - c_j^T x^l), 0)$, where $x^1, \dots, x^l, \dots, x^L$ are all Pareto optimal solutions of (1).

We give a small example to illustrate the method.

$$\begin{aligned} &\min(x_1, x_2) \\ \text{subject to } &2x_1 + 3x_2 \geq 11 \\ &x_1, x_2 \leq 4 \\ &x_1, x_2 \geq 0, \text{ integer} \end{aligned}$$

This problem has Pareto optimal solutions (0,4), (1,3), (3,2), and (4,1), of which (3,2) is unsupported, i.e. not optimal for (2) for any choice of λ. To identify (3,2) as an optimal solution of an elastic constraint problem (4), we choose $i = 2, \varepsilon_1 = 3$ to get

$$\begin{aligned} &\min_{x \in X} x_2 + p_1 su_1 \\ \text{subject to } &x_1 + sl_1 - su_1 = 3 \\ &x \in X. \end{aligned}$$

It is easy to see that for all $p_1 > 1$ (3,2) is a unique optimal solution of the problem. Indeed, there is a whole range of parameters that yield (3,2) as optimal solution.

It is interesting to note two special cases. First consider $\varepsilon_j = \min\{c_j^T x : x \in X\}$, $j \neq i$. This implies that all slack variables sl_j are zero and thus the scalarized problem is equivalent to a weighted sum problem. Secondly, consider $p_j = \infty$, $j \neq i$. Here, all surplus variables su_j must be zero and the problem is equivalent to an ε-constraint problem. The method of elastic constraints is therefore a common generalization of the weighted sum and the ε-constraint method. In contrast to other such methods with this property, e.g. [5], our scalarized problem does not retain the "hard" form of the constraints present in (3), which turns out to be its major advantage, as will be shown below. Further discussion of the method can be found in [3].

3 Bicriteria Airline Crew Scheduling: Cost and Robustness

Airline crew scheduling consists of two distinct problems. The tour of duty (ToD) or pairings planning problem, which involves the construction of tours of duty, and the rostering problem, which means the allocation of ToDs to individual crew members. Here we consider the ToD planning problem. A tour of duty consists of sequences of flights and rest periods that can be operated by a crew member. This problem can be formulated as a generalized set partitioning model, in which the variables are all legal ToDs and the constraints guarantee that each flight in the flight schedule is covered by exactly one ToD. Additional base constraints with non-unit right hand sides are included to take into account the available crew at each base. The objective is to minimize cost. These problems are solved using LP relaxations and branch and bound. Column generation techniques are used to accommodate the huge number of possible variables. The branching process is specifically designed for the model using so-called constraint branching. An overview on optimization techniques for airline crew scheduling can be found in [1].

Statistical delay information from airline operation indicates that minimal cost sets of ToDs can contribute to increasing delays throughout the day because crew are sometimes required to change aircraft. If in that case insufficient ground time is scheduled a delay of the aircraft on which the crew member arrives might cause a delay for the aircraft on which that crew member departs. Our goal was to find solutions of the ToD planning problem which do avoid this behaviour as far as possible while at the same time remaining cost effective. We developed a linear objective function that measures the vulnerability of a solution to disruptions (i.e. the non-robustness of a set of ToDs, see [4]). This yields a bicriteria tour of duty planning model:

$$\begin{aligned} \min[r^T x, c^T x] \\ \text{subject to } A_1 x &= e \\ A_2 x &= b \\ x &\in \{0,1\}. \end{aligned}$$

Here $A_1 x = e$, where $e = (1, \dots, 1)^T$ are the flight constraints (each flight is in exactly one ToD) and $A_2 x = b$ are the so-called base constraints mentioned earlier. The elastic constraint reformulation of the problem is as follows.

$$\begin{aligned} \min r^T x + psu \\ \text{subject to } c^T x + sl - su &= \varepsilon \\ A_1 x &= e \\ A_2 x &= b \\ x &\in \{0,1\}, \end{aligned}$$

where ε is a desired cost level that can be given as a certain percentage increase over the optimal integer solution of the single objective problem of minimizing cost alone.

4 Numerical Results

We implemented the weighted sum method, the ε-constraint method and the method of elastic constraints and solved real world ToD planning problems for technical (pilot and first officer) and cabin crew.

The results show that the weighted sum method finds too few solutions, and also that the intervals of the parameter λ that yield some of these solutions when the weighted sum objective $\min c^T x + \lambda r^T x$ is used are very small. This shows that the solution is very sensitive to small changes of λ. It is therefore not useful in practice. The ε-constraint method needed unacceptable computation times, sometimes exceeding the node limit of 1000 nodes before finding an optimal solution.

For the elastic constraint method the computation times did strongly depend on the value of p: The smaller p, the quicker the problem was solved. Note that larger values of p yield problems that approach the ε-constraint scalarization with its computational difficulties.

Below we show the solutions found on two problem instances for technical crew (Fig. 1) and cabin crew (Fig. 2). Here we increased the value of ε in steps of 0.5% over the cost optimal solution until the optimal value of robustness was reached. The penalty was chosen to be the slope of the efficient frontier of the LP relaxation (the piecewise linear curve shown in Figs. 1 and 2) at the chosen right hand side value of the cost constraint and bound gaps of 0%, 2%, and 20% were used. In both figures the vertical line indicates the optimal cost when solving the single objective problem. The non-robustness objective coefficient r_j for a ToD was computed as a sum of a penalty value that reflects the non-robustness caused by any pair of subsequent flights in a ToD. This penalty is either 0 if both flights are on the same aircraft or the ground duty time plus the expected delay of the incoming flight plus two times the variance of that delay minus the scheduled ground time between the flights, if that number is positive. More detailed results are given in [4].

5 Conclusion

We have shown that well known methods for solving multiobjective optimization problems fail to solve large scale multiobjective combinatorial optimization problems that appear in real world applications. As a consequence we developed the method of elastic constraints which combines the advantages of the weighted sum and ε-constraint methods but avoids their disadvantages. The method was used in a real world application to solve airline crew scheduling problems. The numerical results illustrate the potential of the method of elastic constraints to solve MOCO problems of practically relevant sizes, i.e. large scale problems.

Fig. 1. Set of solutions for technical crew.

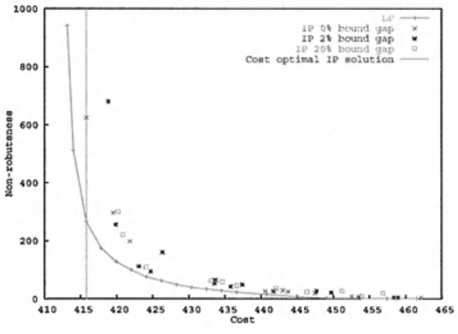
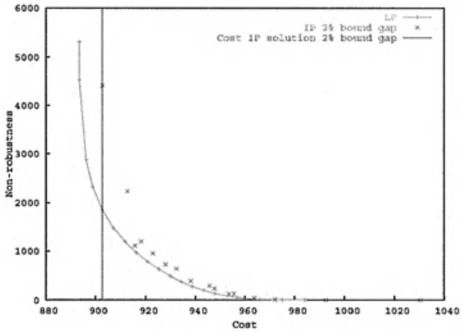


Fig. 2. Set of solutions for cabin crew.



References

1. E.R. Butchers, P.R. Day, A.P. Goldie, S. Miller, J.A. Meyer, D.M. Ryan, A.C. Scott, and C.A. Wallace. Optimised crew scheduling at Air New Zealand. *Interfaces*, 31(1):30–56, 2001.
2. M. Ehrgott and X. Gandibleux. A survey and annotated bibliography of multiobjective combinatorial optimization. *Operations Research Spektrum*, 22:425–460, 2000.
3. M Ehrgott and D.M. Ryan. Multiobjective combinatorial optimization: Scalarizations, relaxations, and the method of elastic constraints. Technical report, Department of Engineering Science, University of Auckland, 2002.
4. M. Ehrgott and D.M. Ryan. Constructing robust crew schedules with bicriteria optimization. *Journal of Multi-Criteria Decision Analysis*, 2003, in print.
5. R.M. Soland. Multicriteria optimization: A general characterization of efficient solutions. *Decision Sciences*, 10:26–38, 1979.

Some Evaluations Based on DEA with Interval Data

Tomoe Entani[1] and Hideo Tanaka[2]

[1] Kochi University, 2-5-1, Akebono-cho, Kochi 780-8520, JAPAN
[2] Hiroshima International University, Gakuendai 555-36, Kurose-cho, Hiroshima 724-0695, JAPAN

Abstract. We propose a new approach for obtaining two interval efficiency values with interval data as an extension of DEA. We deal with interval data that can reflect uncertainty in real situations. The two interval efficiency values are obtained from the optimistic and pessimistic viewpoints. Their upper and lower bounds are obtained by two different extreme values in the given interval data respectively. Thus, we formulate four types of efficiency values from two viewpoints with two extreme values in the given interval data. Our emphasis is to obtain two interval efficiency values reflecting uncertainty of the given data. Thus our approach can be described as a kind of interval data analysis. A numerical example is shown to illustrate our proposed approach.

1 Introduction

DEA (Data Envelopment Analysis) is a non-parametric technique for measuring the efficiency of DMUs which stand for Decision Making Units with common input and output terms [1]. In DEA, the ratio of weighted sum of output data to that of input data is assumed to be the efficiency of the DMU. The input and output weights are variables that are determined so as to maximize the analyzed DMU's ratio subject to the condition concerning every DMU. Therefore, the efficiency value obtained by the optimal weights is regarded as the relative evaluation value from the optimistic viewpoint. This is so-called DEA generally.

We have already formulated the efficiency value from the pessimistic viewpoint in [4] where the input and output weights are determined to minimize the ratio of the analyzed DMU to the other DMUs. Thus, the efficiency is estimated as an interval value constituted of the optimistic and pessimistic efficiency values.

In this paper, we deal with interval data that reflect uncertainty in real situations. Since there is no doubt that real problems contain some uncertainty, uncertain phenomena need to be handled. Various studies [2][5] have been discussed on uncertainty in real problems. For instance monthly sales change every month depending on such as economic situations, seasons and so on. Thus monthly sales can be regarded as an interval value including all the given possible values. We have to deal with interval data reflecting uncertainty in real problems.

The aim of this study is to obtain two interval efficiency values. They are formulated by four types of efficiency values as [max min, max max] and [min min, min max] where [ ] denotes an interval value. Each of the inner operations, max or min, is based on the optimistic or pessimistic viewpoint respectively. Each of the outer operations, min or max, is based on the extreme values in the given interval data. The former and latter interval efficiency values are called the optimistic and pessimistic approximations of efficiency. We put our emphasis on interval efficiency values from two different viewpoints. Our results obtained by our proposed approach might be adequate for real situations.

Lastly, a numerical example is shown to illustrate the proposed approach.

2 Relative Efficiency Value

The relative efficiency value of the analyzed unit DMU_o is denoted as follows.

$$\theta_o = \frac{\frac{\boldsymbol{u}^t\boldsymbol{y}_o}{\boldsymbol{v}^t\boldsymbol{x}_o}}{\max_j\left(\frac{\boldsymbol{u}^t\boldsymbol{y}_j}{\boldsymbol{v}^t\boldsymbol{x}_j}\right)} \tag{1}$$

where $\boldsymbol{x}_j$ and $\boldsymbol{y}_j$ are the given input and output vector of DMU_j whose elements are all positive, and $\boldsymbol{v}$ and $\boldsymbol{u}$ are the weight variables. The numbers of input and output data and DMUs are m, k and n, respectively. The ratio of weighted sum of output data to that of input data for DMU_o is compared to the maximum ratio of all DMUs.

By maximizing or minimizing (1) with respect to the weight variables, $\boldsymbol{v}$ and $\boldsymbol{u}$, θ_o is approximated by two kind of values, θ_o^* and θ_{o*}. They are the extreme values of the relative efficiency values from the optimistic and pessimistic viewpoints for DMU_o.

The problem to obtain θ_o^* is formulated as follows.

$$\begin{aligned} \theta_o^* = \max_{\boldsymbol{u},\boldsymbol{v}} & \frac{\frac{\boldsymbol{u}^t\boldsymbol{y}_o}{\boldsymbol{v}^t\boldsymbol{x}_o}}{\max_j\left(\frac{\boldsymbol{u}^t\boldsymbol{y}_j}{\boldsymbol{v}^t\boldsymbol{x}_j}\right)} \\ \text{s.t. } & \boldsymbol{u} \geq \boldsymbol{0}, \quad \boldsymbol{v} \geq \boldsymbol{0} \end{aligned} \tag{2}$$

The weight variables are determined to maximize the relative efficiency value θ_o. Then we call θ_o^* optimistic approximation of θ_o.

On the other hand, by minimizing θ_o with respect to the weight vriables, θ_{o*} is obtained by the following problem.

$$\begin{aligned} \theta_{o*} = \min_{\boldsymbol{u},\boldsymbol{v}} & \frac{\frac{\boldsymbol{u}^t\boldsymbol{y}_o}{\boldsymbol{v}^t\boldsymbol{x}_o}}{\max_j\left(\frac{\boldsymbol{u}^t\boldsymbol{y}_j}{\boldsymbol{v}^t\boldsymbol{x}_j}\right)} \\ \text{s.t. } & \boldsymbol{u} \geq \boldsymbol{0}, \quad \boldsymbol{v} \geq \boldsymbol{0} \end{aligned} \tag{3}$$

θ_{o*} is called pessimistic approximation of θ_o, since the relative ratio θ_o is minimized.

The relative efficiency value and two approximations by (2) and (3) have the inequality relation $\theta_{o*} \leq \theta_o \leq \theta_o^*$. The efficinecy value measured relatively is smaller than the optimistic approximation and greater than the pessimistic one. We use two extreme values obtained from the optimisitc and pessimistic viewpoints to approximate and denote the relative efficiency value. In the previous study [4], the relative efficiency value is defined as an interval value $[\theta_{o*}, \theta_o^*]$ based on the possibility concept.

3 Approximations of Relative Efficiency Value with Interval Data

In real situations, there are cases that uncertain phenomena need to be handled. Considering the possibility of all the observations, the input and output data are given as interval values. The interval input and output data are denoted as follows with their upper* and lower$_*$ bounds.

input data	ouput data
$x_{jr} \in [x_{jr*}, x_{jr}^*]$	$y_{jp} \in [y_{jp*}, y_{jp}^*]$
$\boldsymbol{x}_j^* = (x_{j1}^*, ..., x_{jm}^*)^t$	$\boldsymbol{y}_j^* = (y_{j1}^*, ..., y_{jk}^*)^t$
$\boldsymbol{x}_{j*} = (x_{j1*}, ..., x_{jm*})^t$	$\boldsymbol{y}_{j*} = (y_{j1*}, ..., y_{jk*})^t$

These interval values are determined so as to include uncertainty of the given data based on the possibility concept. Using the given interval input and output data, the problems to obtain the extreme values that approximate the relative efficiency value are discussed in this section.

3.1 Optimistic Approximation

The optimistic approximation of the relative efficiency value is obtained as the similar formulation to (2). The given data $\boldsymbol{x}_j$ and $\boldsymbol{y}_j$ in (2) are extended to interval values. Therefore the optimistic approximation Θ_o^* is also obtained as interval value as $\Theta_o^* = [\underline{\theta}_o^*, \overline{\theta}_o^*]$ defined below. When the given data are extended to interval values, the optimistic approximation that is the extreme value from the optimistic viewpoint is also extended to an interval value.

The problem to obtain the upper bound of the interval optimistic approximation Θ_o^* can be defined as

$$\overline{\theta}_o^* = \max_{\boldsymbol{u},\boldsymbol{v}} \max_{\boldsymbol{x}_j,\boldsymbol{y}_j} \frac{\dfrac{\boldsymbol{u}^t\boldsymbol{y}_o}{\boldsymbol{v}^t\boldsymbol{x}_o}}{\max_j \left(\dfrac{\boldsymbol{u}^t\boldsymbol{y}_j}{\boldsymbol{v}^t\boldsymbol{x}_j}\right)} \tag{4}$$
$$\text{s.t. } \boldsymbol{x}_{j*} \leq \boldsymbol{x}_j \leq \boldsymbol{x}_j^*, \quad \boldsymbol{y}_{j*} \leq \boldsymbol{y}_j \leq \boldsymbol{y}_j^*, \quad \boldsymbol{u} \geq \boldsymbol{0}, \quad \boldsymbol{v} \geq \boldsymbol{0}.$$

When the data are the optimistic ones for DMU_o, which means that the input data are x_{o*} and $x_j^*(j \neq o)$ and the outuput data are y_o^* and $y_{j*}(j \neq o)$, the relative ratio is maximized. Instead of maximizing the relative ratio with respect to the given interval data, using these chosen input and otuput data (4) can be written as

$$\begin{aligned} \overline{\theta}_o^* &= \max_{\boldsymbol{u},\boldsymbol{v}} \frac{\frac{\boldsymbol{u}^t \boldsymbol{y}_o^*}{\boldsymbol{v}^t \boldsymbol{x}_{o*}}}{\max_j \left(\frac{\boldsymbol{u}^t \boldsymbol{y}_{j*}}{\boldsymbol{v}^t \boldsymbol{x}_j^*} \right)} \\ &\text{s.t. } \boldsymbol{u} \geq \boldsymbol{0}, \ \ \boldsymbol{v} \geq \boldsymbol{0}. \end{aligned} \tag{5}$$

The lower bound of the interval optimistic approximation Θ_o^* is obtained as follows.

$$\begin{aligned} \underline{\theta}_o^* &= \max_{\boldsymbol{u},\boldsymbol{v}} \min_{\boldsymbol{x}_j,\boldsymbol{y}_j} \frac{\frac{\boldsymbol{u}^t \boldsymbol{y}_o}{\boldsymbol{v}^t \boldsymbol{x}_o}}{\max_j \left(\frac{\boldsymbol{u}^t \boldsymbol{y}_j}{\boldsymbol{v}^t \boldsymbol{x}_j} \right)} \\ &\text{s.t. } \boldsymbol{x}_{j*} \leq \boldsymbol{x}_j \leq \boldsymbol{x}_j^*, \ \ \boldsymbol{y}_{j*} \leq \boldsymbol{y}_j \leq \boldsymbol{y}_j^*, \quad \boldsymbol{u} \geq \boldsymbol{0}, \ \ \boldsymbol{v} \geq \boldsymbol{0} \end{aligned} \tag{6}$$

Instead of minimizing the relative ratio with respect to the given interval data, the pessimistic data for DMU_o are chosen in the given interval data. With the chosen input data, x_o^* and $x_{j*}(j \neq o)$, and the chosen output data, y_{o*} and $y_j^*(j \neq o)$, (6) can be written as

$$\begin{aligned} \underline{\theta}_o^* &= \max_{\boldsymbol{u},\boldsymbol{v}} \frac{\frac{\boldsymbol{u}^t \boldsymbol{y}_{o*}}{\boldsymbol{v}^t \boldsymbol{x}_o^*}}{\max_j \left(\frac{\boldsymbol{u}^t \boldsymbol{y}_j^*}{\boldsymbol{v}^t \boldsymbol{x}_{j*}} \right)} \\ &\text{s.t. } \boldsymbol{u} \geq \boldsymbol{0}, \ \ \boldsymbol{v} \geq \boldsymbol{0}. \end{aligned} \tag{7}$$

3.2 Pessimistic Approximation

The pessimistic approximation of the relative efficiency value $\Theta_{o*} = [\underline{\theta}_{o*}, \overline{\theta}_{o*}]$ is formulated in the same way as in (3) with interval input and output data, $\boldsymbol{x}_j$ and $\boldsymbol{y}_j$. The problem to obtain the upper bound of the interval pessimistic approximation Θ_{o*} is as follows.

$$\begin{aligned} \overline{\theta}_{o*} &= \min_{\boldsymbol{u},\boldsymbol{v}} \max_{\boldsymbol{x}_j,\boldsymbol{y}_j} \frac{\frac{\boldsymbol{u}^t \boldsymbol{y}_o}{\boldsymbol{v}^t \boldsymbol{x}_o}}{\max_j \left(\frac{\boldsymbol{u}^t \boldsymbol{y}_j}{\boldsymbol{v}^t \boldsymbol{x}_j} \right)} \\ &\text{s.t. } \boldsymbol{x}_{j*} \leq \boldsymbol{x}_j \leq \boldsymbol{x}_j^*, \ \ \boldsymbol{y}_{j*} \leq \boldsymbol{y}_j \leq \boldsymbol{y}_j^*, \quad \boldsymbol{u} \geq \boldsymbol{0}, \ \ \boldsymbol{v} \geq \boldsymbol{0} \end{aligned} \tag{8}$$

The optimistic data for DMU_o that are chosen in the same way as in the optimistic approximation model (5) are used. Without taking the max operation with respect to the given interval data, (8) can be written as

$$\begin{aligned} \overline{\theta}_{o*} &= \min_{\boldsymbol{u},\boldsymbol{v}} \frac{\frac{\boldsymbol{u}^t \boldsymbol{y}_o^*}{\boldsymbol{v}^t \boldsymbol{x}_{o*}}}{\max_j \left(\frac{\boldsymbol{u}^t \boldsymbol{y}_{j*}}{\boldsymbol{v}^t \boldsymbol{x}_j^*} \right)} \\ &\text{s.t. } \boldsymbol{u} \geq \boldsymbol{0}, \ \ \boldsymbol{v} \geq \boldsymbol{0}. \end{aligned} \tag{9}$$

Although in the optimistic approximation model (5) the ratio is maximized with respect to the weight variables, in the pessimistic approximation model (9) it is minimized. Then the relation of upper bounds of both approximations, $\overline{\theta}_{o*} \leq \overline{\theta}^*_o$, is easily found from (5) and (9).

The problem to obtain the lower bound of the interval pessimistic approximation Θ_{o*} is as follows.

$$\underline{\theta}_{o*} = \min_{\boldsymbol{u},\boldsymbol{v}} \min_{\boldsymbol{x}_j,\boldsymbol{y}_j} \frac{\frac{\boldsymbol{u}^t\boldsymbol{y}_o}{\boldsymbol{v}^t\boldsymbol{x}_o}}{\max_j \left(\frac{\boldsymbol{u}^t\boldsymbol{y}_j}{\boldsymbol{v}^t\boldsymbol{x}_j}\right)} \qquad (10)$$
$$\text{s.t. } \boldsymbol{x}_{j*} \leq \boldsymbol{x}_j \leq \boldsymbol{x}^*_j, \quad \boldsymbol{y}_{j*} \leq \boldsymbol{y}_j \leq \boldsymbol{y}^*_j, \quad \boldsymbol{u} \geq \boldsymbol{0}, \quad \boldsymbol{v} \geq \boldsymbol{0}$$

The chosen data are the same as ones in the optimistic approximation model (7). With the pessimistic data for DMU_o (10) can be written as

$$\underline{\theta}_{o*} = \min_{\boldsymbol{u},\boldsymbol{v}} \frac{\frac{\boldsymbol{u}^t\boldsymbol{y}_{o*}}{\boldsymbol{v}^t\boldsymbol{x}^*_o}}{\max_j \left(\frac{\boldsymbol{u}^t\boldsymbol{y}^*_j}{\boldsymbol{v}^t\boldsymbol{x}_{j*}}\right)} \qquad (11)$$
$$\text{s.t. } \boldsymbol{u} \geq \boldsymbol{0}, \quad \boldsymbol{v} \geq \boldsymbol{0}.$$

From (7) and (11), the lower bounds of both approximations have the relation $\underline{\theta}_{o*} \leq \underline{\theta}^*_o$. Therefore, the interval optimistic and pessimistic approximations satisfy the inequality relation $\Theta_{o*} = [\underline{\theta}_{o*}, \overline{\theta}_{o*}] \leq \Theta^*_o = [\underline{\theta}^*_o, \overline{\theta}^*_o]$. The pessimistic approximation is smaller than the optimistic one in a sense of interval order relation in [3], where $A = [\underline{a}, \overline{a}] \leq B = [\underline{b}, \overline{b}]$ holds if and only if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$.

In rough set theory[6], X is approximated with the upper and lower approximations, $A^*(X)$ and $A_*(X)$, that satisfy the inclusion relation $A_*(X) \subset A^*(X)$. The lower approximation is included in the upper one. The inclusion relation in rough set theory corresponds to the inequality relation in this study.

4 Numerical Example

All the given input data are normalized to 1 and two kinds of output data are interval values including uncertainty of the given data in Table 1. Using the data we calculate the interval optimistic approximation by (5) and (7) and the interval pessimistic approximation by (9) and (11).

5 Conclusion

The relative efficiency value is approximated by two extreme values from the optimistic and pessimisitic viewpoints with optimistic and pessimisitic data for the analyzed unit DMU_o. The pessimistic approximation is always smaller

Table 1. Interval data and interval approximations

	input	output1	output2	optimistic approximation	pessimistic approximation
A	1	[0.8,1.2]	[7.5,8.5]	[1.000, 1.000]	[0.110, 0.179]
B	1	[1.8,2.2]	[2.4,3.6]	[0.416, 0.634]	[0.247, 0.328]
C	1	[1.7,2.3]	[5.7,6.3]	[0.729, 0.923]	[0.233, 0.343]
D	1	[2.5,3.5]	[2.7,3.3]	[0.525, 0.786]	[0.318, 0.440]
E	1	[2.8,3.2]	[6.7,7.3]	[0.964, 1.000]	[0.384, 0.477]
F	1	[3.8,4.2]	[1.8,2.2]	[0.614, 0.782]	[0.212, 0.293]
G	1	[3.4,4.6]	[4.6,5.4]	[0.801, 1.000]	[0.466, 0.687]
H	1	[4.7,5.3]	[1.5,2.5]	[0.702, 0.962]	[0.176, 0.333]
I	1	[5.6,6.4]	[1.7,2.3]	[0.829, 1.000]	[0.200, 0.307]
J	1	[6.7,7.3]	[0.8,1.2]	[1.000, 1.000]	[0.094, 0.160]

than the optimistic one. The case that the input and output data are given as interval values has been dealt with in this paper. Interval data including all the observations are suitable to reflect uncertainty in real situations. Then the interval optimistic and pessimistic approximations of a DMU are obtained to reflect uncertainty in real situations.

What we have done with the given interval data is to denote the relative efficiency values of DMUs by their interval optimistic and pessimistic approximations. The interval efficiency values are useful information for the analyzed DMUs. The order relation over DMUs with them is not discussed in this paper, but it might depend on a decision maker's attitude toward the problem.

References

1. Charnes, A., Cooper, W. W. and Rhodes, E. (1978) Measuring the Efficiency of Decision Making Units, European Journal of Operational Research, **2**, 429–444.
2. Despotis, D. K. and Smirlis, Y.G. (2002) Data Envelopment Analysis with Imprecise Data, European Journal of Operational Research, **140**, 24–36.
3. Dubious, D. and Prade H. (1980) Systems of Linear Fuzzy Constraints, Fuzzy Sets and Systems, **3**, 37–48.
4. Entani, T., Maeda, Y. and Tanaka, H. (2002) Dual Models of Interval DEA and Its Extension to Interval Data, European Journal of Operational Research, **136**, 32–45.
5. Inuiguchi, M. and Tanino, T. (2000) Data Envelopment Analysis with Fuzzy Input-output Data, in: Haimes, Y. Y. and Steuer, R. E. (Eds.): Research and Practice in Multiple Criteria Decision Making, Springer-Verlag, Berlin, 296–307.
6. Pawlak, Z. (1984) Rough Classification, Journal of Man-Machine Studies, **20**, 469–482.

Possibility and Necessity Measures in Dominance-based Rough Set Approach

Salvatore Greco[1], Masahiro Inuiguchi[2], and Roman Słowiński[3]

[1] Faculty of Economics, University of Catania, Corso Italia, 55, 95129 Catania, Italy, salgreco@unict.it
[2] Graduate School of Engineering, Osaka University, Suita, Osaka 565-0871, Japan, inuiguti@eie.eng.osaka-u.ac.jp
[3] Institute of Computing Science, Poznan University of Technology, 60-965 Poznan, Poland, slowinsk@sol.put.poznan.pl

Abstract. Dominance-based rough set approach is an extension of the basic rough set approach proposed by Pawlak, to multicriteria classification problems. In this paper, the dominance-based rough set approach is considered in the context of vague information on preferences and decision classes. The vagueness is handled by possibility and necessity measures defined using modifiers of fuzzy sets. Due to this way of handling the vagueness, the lower and upper approximations of preference-ordered decision classes are fuzzy sets whose membership functions are necessity and possibility measures, respectively.

1 Introduction

The rough set theory has been proposed by Z. Pawlak [5] to deal with inconsistency problems following from information granulation. The original rough set idea has proved to be particularly useful in the analysis of multiattribute classification problems; however, it was failing when attributes whose domains are preference-ordered (criteria) had to be taken into account. Indeed, in many real problems it is important to consider the ordinal properties of the considered criteria. For example, in bankruptcy risk evaluation, if the debt index (total debt/total activity) of firm A has a modest value, while the same index of firm B has a significant value, then, within the rough set approach, the two firms are merely discernible, but no preference is given to one of them with reference to the attribute "debt ratio". In reality, from the point of view of the bankruptcy risk evaluation, it would be advisable to consider firm A better than firm B, and not simply different (discernible). Therefore, the attribute "debt ratio" is a criterion. Consideration of criteria in rough set approximation can be made by replacing indiscernibility or similarity relation by the dominance relation, which is a very natural concept within multicriteria decision making.

In order to deal with problems of multicriteria decision making (MCDM), like sorting, choice or ranking, a number of methodological changes to the original rough set theory were necessary [2]. The main change is the substitution of the indiscernibility relation by a dominance relation (crisp or fuzzy),

Table 1. An example of Q_h^i, $i = 1, 2$, m_h and M_h

Modifier	0	$\rightarrow$	h	$\rightarrow$	1
Q_h^1	most weakly	more or less	normally	very	most strongly
Q_h^2	most weakly	more or less	normally	very	most strongly
m_h	most strongly	very	normally	more or less	most weakly
M_h	most weakly	more or less	normally	very	most strongly

which permits approximation of ordered sets in multicriteria sorting. In this paper we propose a fuzzy extension of the rough approximation by dominance relation based on the concepts of necessity and possibility. In particular, we are considering a special definition of necessity and possibility measures introduced in [4]. For an alternative fuzzy extension of rough approximation by dominance relation see [2,3]. The paper is organized as follows. Section two recalls necessity and possibility measures. Section three presents basic idea of rough approximation by fuzzy dominance. Conclusions are grouped in the last section.

2 Possibility and Necessity Measures

Possibility and necessity measures are defined by

$$\Pi(B|A) = \sup_x C(\mu_A(x), \mu_B(x)), \qquad N(B|A) = \inf_x I(\mu_A(x), \mu_B(x)),$$

where A and B are fuzzy sets with membership functions μ_A and μ_B. C and $I : [0,1] \times [0,1] \rightarrow [0,1]$ are conjunction and implication functions such that

C1) $C(0,0) = C(0,1) = C(1,0) = 0$ and $C(1,1) = 1$,
I1) $I(0,0) = I(0,1) = I(1,1) = 1$ and $I(1,0) = 0$.

We often use monotonic conjunction and implication functions C and I which satisfy

C2) $C(a,b) \leq C(c,d)$ if $a \leq c$ and $b \leq d$,
I2) $I(a,b) \leq I(c,d)$ if $a \geq c$ and $c \leq d$.

If C and I satisfies C2) and I2) then we have the following properties, respectively:

$$\Pi(B_1|A_1) \leq \Pi(B_2|A_2) \text{ and } N(B_1|A_2) \leq N(B_2|A_1) \text{ if } A_1 \subseteq A_2 \text{ and } B_1 \subseteq B_2.$$

In the following we use also a negation function $neg : [0,1] \rightarrow [0,1]$, such that $neg(0) = 1$, $neg(neg(a)) = a$ and neg is a non-increasing function.

Since there exist many conjunction and implication functions, we have also many possibility and necessity measures. Thus there is a question, how we select possibility and necessity measures. To answer this question, in [4] the level cut conditioning approach has been proposed. In this approach, we can specify possibility and necessity measures based on the following equivalences:

$$\Pi(B|A) \leq h \text{ if and only if } Q_h^1(A) \subseteq \left(Q_h^2(B)\right)^{\mathrm{c}}, \tag{1}$$

$$N(B|A) \geq h \text{ if and only if } m_h(A) \subseteq M_h(B), \tag{2}$$

where A^{c} is the complement fuzzy set of A and inclusion relation $A \subseteq B$ is defined by $\mu_A \leq \mu_B$. Q_h^i, $i = 1, 2$, m_h and M_h are modifiers varying with a parameter $h \in (0, 1)$. An example of Q_h^i, $i = 1, 2$, m_h and M_h is given in Table 1. As h becomes large, condition $Q_h^1(A) \subseteq \left(Q_h^2(B)\right)^{\mathrm{c}}$ becomes weak while condition $m_h(A) \subseteq M_h(B)$ becomes strong.

In order to proceed with calculations, $Q_h^i(A)$, $i = 1, 2$, $m_h(A)$ and $M_h(A)$ are defined by the following membership functions:

$$\mu_{Q_h^i(A)}(x) = g_i^Q(\mu_A(x), h),\ \ i = 1, 2,$$
$$\mu_{m_h(A)}(x) = g^m(\mu_A(x), h) \text{ and } \mu_{M_h(A)}(x) = g^M(\mu_A(x), h).$$

From the properties of modifiers Q_h^i, m_h and M_h, modifier functions g_i^Q, g^m and g^M should satisfy the following requirements: q1) $g_i^Q(a, \cdot)$ is lower semi-continuous for all $a \in [0, 1]$, q2) $g_i^Q(1, h) = 1$ and $g_i^Q(0, h) = 0$ for all $h < 1$, q3) $g_i^Q(a, 1) = 0$ for all $a \in [0, 1]$, q4) $g_i^Q(a, \cdot)$ is non-increasing for all $a \in [0, 1]$, q5) $g_i^Q(\cdot, h)$ is non-decreasing for all $h \in [0, 1]$, q6) $g_i^Q(a, 0) > 0$ for all $a \in (0, 1)$, g1) $g^m(a, \cdot)$ and $g^M(a, \cdot)$ are lower and upper semi-continuous for all $a \in [0, 1]$, respectively, g2) $g^m(1, h) = g^M(1, h) = 1$ and $g^m(0, h) = g^M(0, h) = 0$ for all $h > 0$, g3) $g^m(a, 0) = 0$ and $g^M(a, 0) = 1$ for all $a \in [0, 1]$, g4) $g^m(a, \cdot)$ is non-decreasing and $g^M(a, \cdot)$ is non-increasing for all $a \in [0, 1]$, g5) $g^m(\cdot, h)$ and $g^M(\cdot, h)$ are non-decreasing for all $h \in [0, 1]$ and g6) $g^m(a, 1) > 0$ and $g^M(a, 1) < 1$ for all $a \in (0, 1)$.

Given modifier functions g_i^Q $(i = 1, 2)$, g^m and g^M, it is shown that possibility and necessity measures are obtained as (see [4])

$$\Pi^L(B|A) = \inf_h \{h \in [0, 1] \mid Q_h^1(A) \subseteq (Q_h^2(B))^{\mathrm{c}}\} = \sup_x C^L(\mu_A(x), \mu_B(x)),$$
$$N^L(B|A) = \sup_h \{h \in [0, 1] \mid m_h(A) \subseteq M_h(B)\} = \inf_x I^L(\mu_A(x), \mu_B(x)),$$

where conjunction function C^L and implication function I^L are defined by

$$C^L(a, b) = \inf_h \{h \in [0, 1] \mid g_1^Q(a, h) \leq neg(g_2^Q(b, h))\},$$
$$I^L(a, b) = \sup_h \{h \in [0, 1] \mid g^m(a, h) \leq g^M(b, h)\}.$$

Conjunction and implication functions C^L and I^L satisfy C2) and I2), respectively. It is shown that many famous conjunction functions and implication functions are obtained from modifier functions g_i^Q (i=1,2), g^m and g^M. Properties of C^L and I^L are investigated in [4].

3 Approximations by Means of Fuzzy Dominance Relations

Let us remember that formally, by an *information table* we understand the 4-tuple $S = \langle U, Q, V, f \rangle$, where U is a finite set of objects, Q is a finite set of *attributes*, $V = \bigcup_{q \in Q} V_q$ and V_q is a domain of the attribute q, and $f : U \times Q \to V$ is a total function such that $f(x,q) \in V_q$ for every $q \in Q$, $x \in U$, called an *information function* (cf. [5]).

Furthermore an information table can be seen as *decision table* assuming that the set of attributes $Q = K \cup D$ and $K \cap D = \emptyset$, where set K contains so-called *condition attributes*, and D, *decision attributes*. In the dominance-based rough set approach we are considering attributes with preference-ordered domains – such attributes are called *criteria*.

In this section we refine the concept of dominance-based rough approximation recalled in section 2, by introducing gradedness through the use of fuzzy sets. In the following we shall use the concepts of T-norm T and T-conorm T^* defined as follows: $T : [0,1] \times [0,1] \to [0,1]$ such that for each $a,b,c,d \in [0,1]$, $T(a,b) \geq T(c,d)$ when $a \geq c$ and $b \geq d$, $T(a,1) = a$, $T(a,b) = T(b,a)$ and $T(a, T(b,c)) = T(T(a,b),c)$; $T^* : [0,1] \times [0,1] \to [0,1]$ such that for each $a,b,c,d \in [0,1]$, $T^*(a,b) \geq T^*(c,d)$ when $a \geq c$ and $b \geq d$, $T^*(a,0) = a$, $T^*(a,b) = T^*(b,a)$ and $T^*(a, T^*(b,c)) = T^*(T^*(a,b),c))$.

Let S_q be a fuzzy outranking relation on U with respect to criterion $q \in K$, i.e. $S_q : U \times U \to [0,1]$, such that $S_q(x,y)$ represents the credibility of the proposition "x is at least as good as y with respect to criterion q". It is natural to consider S_q as a fuzzy partial T-preorder, i.e. reflexive (for each $x \in U$, $S_q(x,x) = 1$) and T-transitive (for each $x,y,z \in U$, $T(S_q(x,y), S_q(y,z)) \leq S_q(x,z)$). Fuzzy outranking relation S_q can be build from another fuzzy complete T-preorder defined on domain V_q of criterion $q \in K$, i.e. $S_{V_q} : V_q \times V_q \to [0,1]$ such that $S_q(x,y) = S_{V_q}(f(x,q), f(y,q))$.

Using the fuzzy outranking relations S_q, $q \in K$, a fuzzy dominance relation on U (denotation $D_P(x,y)$) can be defined for each $P \subseteq K$ as follows:

$$D_P(x,y) = \underset{q \in P}{T} S_q(x,y).$$

Given $(x,y) \in U \times U$, $D_P(x,y)$ represents the credibility of the proposition "x outranks y on each criterion q from P". Let us remark that from the reflexivity of fuzzy outranking S_q, $q \in K$, we have that for each $x \in U$ $D_P(x,x) = 1$, i.e. also D_P is reflexive.

Since the fuzzy outranking relations S_q are supposed to be partial T-preorders, then also the fuzzy dominance relation D_P is a partial T-preorder. Furthermore, let $\boldsymbol{Cl} = \{Cl_t, t \in H\}$, $H = \{1, ..., n\}$, be a set of fuzzy classes in U, such that for each $x \in U$, $Cl_t(x)$ represents the membership function of x to Cl_t. We suppose that the classes of $\boldsymbol{Cl}$ are ordered according to increasing preference, i.e. that for each $r, s \in H$, such that $r > s$, the elements of Cl_r have a better comprehensive evaluation than the elements of Cl_s.

For example, in a problem of bankruptcy risk evaluation, Cl_1 is the set of unacceptable risk firms, Cl_2 is a set of high risk firms, Cl_3 is a set of medium risk firms, and so on.

On the basis of the membership functions of the fuzzy class Cl_t, we can define fuzzy membership functions of two merged fuzzy sets:

1) the upward merged fuzzy set $Cl_t^{\geq}$, whose membership function $Cl_t^{\geq}(x)$ represents the credibility of the proposition "x is at least as good as the objects in Cl_t",

$$Cl_t^{\geq}(x) = \begin{cases} 1 & \text{if } \exists s \in H : Cl_s(x) > 0 \text{ and } s > t \\ Cl_t(x) & \text{otherwise,} \end{cases}$$

2) the downward merged fuzzy set $Cl_t^{\leq}$, whose membership function $Cl_t^{\leq}(x)$ represents the credibility of the proposition "x is at most as good as the objects in Cl_t",

$$Cl_t^{\leq}(x) = \begin{cases} 1 & \text{if } \exists s \in H : Cl_s(x) > 0 \text{ and } s < t \\ Cl_t(x) & \text{otherwise.} \end{cases}$$

We say that the credibility of the statement "x belongs without ambiguity to $Cl_t^{\geq}$" is equal to the degree of necessity of the statement "all objects $y \in U$ dominating x belong to $Cl_t^{\geq}$". Furthermore, we say that the credibility of the statement "x possibly belongs to $Cl_t^{\geq}$" is equal to the degree of possibility of the statement "some object $y \in U$ dominated by x belongs to $Cl_t^{\geq}$". Analogous statements can be formulated for inclusion of x in $Cl_t^{\leq}$.

Therefore, the P-lower and the P-upper approximations of $Cl_t^{\geq}$ with respect to $P \subseteq K$ are fuzzy sets in U, whose membership functions (denotation $\underline{P}[Cl_t^{\geq}(x)]$ and $\overline{P}[Cl_t^{\geq}(x)]$, respectively) are defined as:

$$\underline{P}[Cl_t^{\geq}(x)] = N(Cl_t^{\geq}|D_p^{+}(x)) = \inf_{y \in U} I(D_p(y,x), Cl_t^{\geq}(y)),$$
$$\overline{P}[Cl_t^{\geq}(x)] = \Pi(Cl_t^{\geq}|D_p^{-}(x)) = \sup_{y \in U} C(D_p(x,y), Cl_t^{\geq}(y)),$$

where $D_p^{+}(x)$ is a fuzzy set of objects $y \in U$ dominating x with respect to $P \subseteq K$ and $D_p^{-}(x)$ is a fuzzy set of objects $y \in U$ dominated by x with respect to $P \subseteq K$. The membership functions of $D_p^{+}(x)$ and $D_p^{-}(x)$ are defined as:

$$\mu(y, D_p^{+}(x)) = D_p(y,x), \qquad \mu(y, D_p^{-}(x)) = D_p(x,y).$$

The P-lower and P-upper approximations of $Cl_t^{\leq}$ with respect to $P \subseteq K$ (denotation $\underline{P}[Cl_t^{\leq}(x)]$ and $\overline{P}[Cl_t^{\leq}(x)]$) can be defined, analogously.

Greco, Inuiguchi and Słowiński proved that the basic properties of rough set theory hold for the above definitions of lower and upper approximations subject to some conditions [1].

4 Conclusion

We introduced fuzzy rough approximation using fuzzy dominance to deal with multicriteria sorting problems. We proved that our extension of rough approximation maintains the same desirable properties of classical rough set approximation within fuzzy set context. Given the rough approximations of fuzzy decision classes being merged according to the preference order, one is able to induce certain and possible decision rules from these approximations. Each certain rule is characterized by a necessity degree, and each possible rule, by a possibility degree, corresponding to rule credibility. For example, in the context of credit analysis a decision rule can have a syntax like "if the debt ratio is not larger than 3 and the return on investment not smaller than 10%, then the firm is at most a medium risk firm; the credibility of this implication being equal to 75%."

Acknowledgements

The first author wishes to acknowledge financial support from the Italian Ministry of Education, University and Scientific Research (MIUR). The research of the third author has been supported by the State Committee for Scientific Research (KBN), research grant no. 8T11F 006 19, and by the Foundation for Polish Science, subsidy no. 11/2001.

References

1. Greco, S., M. Inuiguchi and R. Słowiński. "Dominance-based Rough Set Approach Using Possibility and Necessity Measures". *Proceedings of the Third International Conference, RSCTC (Rough Sets and Current Trends in Computing) 2002 October 14-16, 2002, Penn State Great Valley, Lecture Notes in Computer Science / Lecture Notes in Artificial Intelligence (LNCS/LNAI)*, Springer-Verlag, (in print)
2. Greco, S., B. Matarazzo and R. Słowiński. "The use of rough sets and fuzzy sets in MCDM". In: T. Gal, T. Hanne and T. Stewart (eds.): *Advances in Multiple Criteria Decision Making.* Kluwer Academic Publishers, Dordrecht, Boston, chapter 14, pp.14.1–14.59, 1999.
3. Greco, S., B. Matarazzo and R. Słowiński. "A fuzzy extension of the rough set approach to multicriteria and multiattribute sorting". In J. Fodor, B. De Baets and P. Perny (eds.): *Preferences and Decisions under Incomplete Information*, Physica-Verlag, Heidelberg, 2000, pp.131–154.
4. Inuiguchi, M., S. Greco, R. Słowiński and T. Tanino. "Possibility and necessity measure specification using modifiers for decision making under fuzziness". *Fuzzy Sets and Systems*, to appear
5. Pawlak, Z. *Rough Sets. Theoretical Aspects of Reasoning about Data.* Kluwer Academic Publishers, Dordrecht, 1991.

Simplex Coding Genetic Algorithm for the Global Optimization of Nonlinear Functions*

Abdel-Rahman Hedar and Masao Fukushima

Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan

Abstract. A new algorithm called Simplex Coding Genetic Algorithm (SCGA) is proposed for solving nonlinear global optimization problems. This algorithm is obtained by hybridizing genetic algorithm and simplex-based local search method called Nelder-Mead method. The efficiency of SCGA is tested on some well known functions. Comparison with other meta-heuristics indicates that the SCGA is promising.

1 Introduction

Global optimization has drawn much attention recently [5,9], because of a very broad spectrum of applications in real-world systems. In this paper, we focus on the case of unconstrained minimization, i.e., the problem is

$$\min_{x \in R^n} f(x),$$

where f is a generally nonconvex, real valued function defined on R^n. Meta-heuristics contribute to a reasonable extent in solving global optimization problems, mainly combinatorial problems [10]. Genetic algorithms (GAs) are one of the most efficient meta-heuristics [7], that have been employed in a wide variety of problems. However, GAs, like other meta-heuristics, suffer from the slow convergence that brings about the high computational cost.

Recently, several new approaches have been developed to furnish meta-heuristics with the ability to simulate the fast convergence of local search methods. Most of these approaches hybridize local search methods with meta-heuristics to obtain more efficient methods with relatively faster convergence. This paper pursues in that direction and proposes a new hybrid method that combines GA with a local search method called Nelder-Mead method [9]. In the combined method, called the simplex coding genetic algorithm (SCGA), we consider the members of the population to be simplices, i.e., each chromosome is a simplex and the gene is a vertex of this simplex. Selection, crossover and mutation procedures are used to improve the initial population. Moreover, Nelder-Mead method is applied to improve the population in the initial

* This research was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan.

stage and every intermediate stage when new children are generated. In the SCGA, we use the linear ranking selection scheme [1] to choose some fit parents to be mated. Then, using a new scheme of a multi-parents crossover, new children are reproduced and a few of them are mutated. Applying Kelley's modification [6] of Nelder-Mead method on the best point visited is the final stage in the SCGA to accelerate the search and to improve this best point.

There have been some attempts to utilize the idea of hybridizing local search methods with GA. Simple hybrid methods use the GA or local search methods to generate the points for the new population and then apply the other technique to improve this new population. Other hybrid methods do some modifications in the GA operations; selection, crossover and mutation using local search methods [11].

The description of the proposed method is given in the next section. Section 3 discusses the experimental results along with the initialization of some parameters and the setting of the control parameters of the proposed method. The conclusion follows the experimental results and makes up Section 4.

2 Description of SCGA

In this section, we describe the proposed method SCGA. The SCGA uses the main functions of the GA; selection, crossover and mutation, on a population of simplices to encourage the exploration process. Moreover, the SCGA tries to improve the initial members and new children by applying a local search method to enhance the exploitation process. This kind of exploration-exploitation procedure is sometimes called "Memetic Algorithm", see [8]. Finally, the SCGA applies an effective local search method on the best point reached by the previous exploration-exploitation procedure. The purpose of this local search is to accelerate the final stages of the GA procedure. This strategy is expected to be effective because the GA has a difficulty in obtaining some required accuracy although the GA may quickly approach the neighborhood of the global minimum.

2.1 Initialization

The SCGA starts with the following initialization procedure:

1. Generate the initial population P_0 that consists of M chromosomes (simplices), i.e., $P_0 = \left\{ S^j : S^j = \{x^{j,i}\}_{i=1}^{n+1} ;\ x^{j,i} \in R^n,\ j = 1, \dots, M \right\}$.
2. Order the vertices of each simplex S^j, $j = 1, 2, \dots, M$, so that
$$f(x^{j,1}) \leq f(x^{j,2}) \leq \cdots \leq f(x^{j,n+1}). \quad (1)$$
3. Apply a small number of iterations of the Nelder-Mead method with each S^j as an initial simplex to improve the chromosomes in the initial population P_0.

4. Order the simplices $S^j = \{x^{j,i}\}_{i=1}^{n+1}$, $j = 1, \ldots, M$ in the improved population P_0 so that

$$f(x^{1,1}) \leq f(x^{2,1}) \leq \cdots \leq f(x^{M,1}). \tag{2}$$

2.2 GA loop

While the stopping conditions are not satisfied, repeat the following procedures; selection, crossover and mutation, and reduction of the population.

Selection. We describe how we select the set $Q \subseteq P$ of the members that will be given the chance to be mated from the current population P. The number of members in each P or Q stays constant but more fit members in P are chosen with higher probability in Q. We use Baker's scheme called "linear ranking selection" [1] to select the new members in Q.

Crossover and mutation. Choose a random number from the unit interval $[0, 1]$ for each chromosome in Q. If this number is less than the predertermined crossover probability p_c, then this chromosome is chosen as a parent. Repeat the following steps until all parents are mating.

1. Select a number n_c from the set $\{2, \ldots, n+1\}$ randomly to determine the number of parents chosen to be mated together.
2. Compute new children $C^i = \{x_c^{i,k}\}_{k=1}^{n+1}$, $i = 1, \ldots, n_c$ by

$$x_c^{i,k} = \overline{x}^k + d\, r^i, \ k = 1, \ldots, n+1, \tag{3}$$

   where r^i, $i = 1, \ldots, n_c$, are random vectors of length less than 1, d is the maximum distance between pairs of parents and $\overline{x}^k$ is the average of the kth vertices of all parents, i.e.,

$$\overline{x}^k = \frac{1}{n_c} \sum_{i=1}^{n_c} x^{i,k}, \ k = 1, \ldots, n+1. \tag{4}$$

   Figure 1 shows an example of crossover in two dimensions. In Figure 1(a), we use Equations (4) to compute the dotted simplex whose vertices are the average of the vertices of the parents S^1, S^2 and S^3. By using Equations (3), we move this dotted simplex randomly inside the circle to create the children C^1, C^2 and C^3, as in Figure 1(b).
3. Choose a random number from the unit interval $[0, 1]$ for each child C^i, $i = 1, \ldots, n_c$. If this number is less than the predertermined mutation probability p_m, then this child is mutated. Let I_m be the index set of those children who are mutated.
4. Apply the following procedure for each child $C^i = \{x_c^{i,k}\}_{k=1}^{n+1}$, $i \in I_m$. Select a number n_i from the set $\{1, 2, \ldots, n+1\}$ randomly to determine the vertex that is reflected as a mutation. Compute the mutated child $\widetilde{C}^i = \{x_m^{i,k}\}_{k=1}^{n+1}$ by

$$\begin{aligned} x_m^{i,k} &= x_c^{i,k}, \ k = 1, \ldots, n_i - 1, n_i + 1, \ldots, n+1, \\ x_m^{i,n_i} &= \overline{x} + u\left(\overline{x} - x_c^{i,n_i}\right), \end{aligned}$$

where u is a random number in the interval $[0.5, 1.5]$ and $\overline{x}$ is the average of vectors $x_c^{i,1}, \ldots, x_c^{i,n_i-1}, x_c^{i,n_i+1}, \ldots, x_c^{i,n+1}$. Replace the child C^i by the mutated one $\widetilde{C}^i$. Figure 1(c) shows an example of mutation in two dimensions, where the mutated simplex consists of the vertices $x_m^{1,1}$, $x_m^{1,2}$ and $x_m^{1,3}$, where the vertex $x_m^{1,2}$ is randomly chosen on the line segment $\overline{p_1p_2}$.

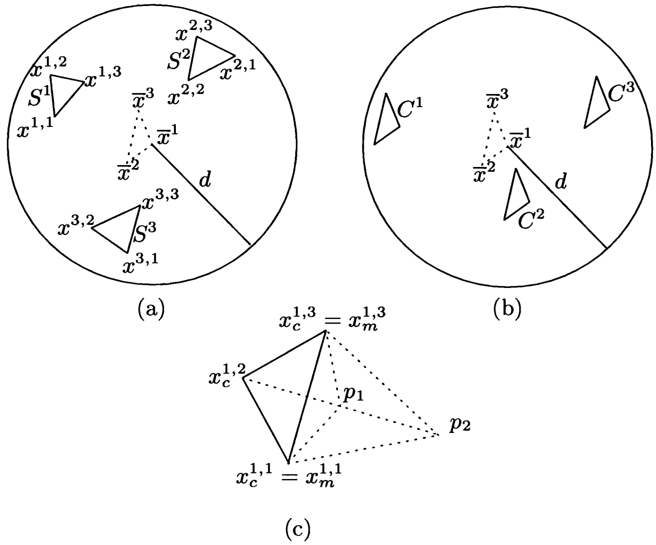


Fig. 1. An example of SCGA crossover and mutation in two dimensions

5. Apply a small number of iterations of the Nelder-Mead method with each child C^i, $i = 1, \ldots, n_c$ as an initial simplex to improve the chromosomes.
6. The population in the next generation consists of the M best ones from the set $P \cup \{C^i\}_{i=1}^{n_c}$. Re-order the chromosomes in the new population so that (1) and (2) hold.

Reduction of the population. After every predetermined number of generations, remove some of the worst members in the population P.
Acceleration in the final stage. From the best point obtained by the above procedures, construct a small simplex. Then, apply Kelley's modification [6] of the Nelder-Mead method on this simplex to obtain the final solution.

3 Experimental Results

GA loop parameters. The steps of the GA loop have been described in the previous section. Here we specify the values of the parameters used in this loop. The control parameter $\eta_{\max}$ in the selection procedure is chosen to be 1.1 according to the original setting in [1]. The crossover probability p_c and the mutation probability p_m are set equal to 0.6 and 0.1, respectively. The number of Nelder-Mead iterations in the local search for the new children is fixed at 2. At every $3n$ generations, we remove the n worst chromosomes from the population unless the number of its chromosomes is less than $2n$.

Termination criteria. The SCGA is terminated when the function values at all vertices of the simplex that contains the best point become close to each other. In order to limit the computations whenever this termination condition cannot be achieved, we terminate the algorithm if the number of generations exceeds the predetermined number set equal to $\min(10n, 100)$.

Numerical results. The performance of the SCGA was tested on a number of well known functions [2,3,4]. For each function we made 100 trials with different initial populations. The SCGA algorithm was programmed in MATLAB and was run on a personal computer running at 733 MHz. To judge the success of a trial, we used the condition: $|f^* - \widehat{f}| < \epsilon_1 |f^*| + \epsilon_2$, where $\widehat{f}$ refers to the best function value obtained by SCGA, f^* refers to the known exact global minimum, and ϵ_1 and ϵ_2 are small positive numbers. We set ϵ_1 and ϵ_2 equal to 10^{-4} and 10^{-6}, respectively. The results are shown in Table 1, where the average number of function evaluations is related to only successful trials. Table 1 shows that the SCGA reached the global minima in a very good successful rate for the majority of the tested functions. In this Table, we also compare the results of the SCGA with those of three other meta-heuristic methods. These methods are Real-value Coding Genetic Algorithm (RCGA) [2], Continuous Genetic Algorithm (CGA) [3] and Direct Search Simulated Annealing (DSSA) [4]. The figures for these methods in Table 1 are taken from the original references. The comparison given in Table 1 shows the DSSA outperforms the other GA methods for the majority of test functions.

4 Conclusion

In this paper, we have introduced a simplex coding genetic algorithm that uses a set of simplices as the population. Applying the Nelder-Mead local search method on these simplices in addition to the ordinary GA operations such as selection, crossover and mutation enhances the exploration process and accelerates the convergence of the GA. We also have introduced a new kind of multi-parents crossover that gives the chance to more than two parents to cooperate in reproducing children and exploring the region around these parents. Moreover, using a local search method again in the final stage helps the GA in obtaining good accuracy quickly. Finally, the computational results show that the SCGA works successfully on some well known test functions.

Table 1. Average number of function evaluations

Function	SCGA	RCGA [2]	CGA [4]	DSSA [8]
Branin	173	490	620	118
Easom	715	642	1504	1442 (93%)
Goldstein	191	270	410	261
Shubert	742 (98%)	946	575	457 (94%)
Michalewicz	179	452	-	-
Rosenbrock (R_2)	222	596	960	306
Zakharov (Z_2)	170	437	620	186
De Joung	187	395	750	273
Hartmann ($H_{3,4}$)	201	342	582	572
Shekel ($S_{4,5}$)	1086 (79%)	1158 (62%)	610 (76%)	993 (81%)
Shekel ($S_{4,7}$)	1087 (81%)	1143 (70%)	680 (83%)	932 (84%)
Shekel ($S_{4,10}$)	1068 (84%)	1235 (58%)	650 (81%)	992 (77%)
Rosenbrock (R_5)	3629 (90%)	4150 (60%)	3990	2685
Zakharov (Z_5)	998	1115	1350	914
Hartmann ($H_{6,4}$)	989 (99%)	973	970	1737 (92%)
Rosenbrock (R_{10})	6340 (90%)	8100 (70%)	21563 (80%)	16785
Zakharov (Z_{10})	1829	2190	6991	12501

References

1. Baker, J. E. (1985) Adaptive selection methods for genetic algorithms. In: Grefenstette, J.J. (Ed.) Proceedings of the First International Conference on Genetic Algorithms, Lawrence Erlbaum Associates, Hillsdale, MA, 101–111
2. Bessaou, M., Siarry, P. (2001) A genetic algorithm with real-value coding to optimize multimodal continuous functions. Struct Multidisc Optim. **23**, 63-74
3. Chelouah, R., Siarry, P. (2000) A continuous genetic algorithm designed for the global optimization of multimodal functions. J. Heuristics. **6,** 191-213
4. Hedar, A., Fukushima, M. (2002) Hybrid simulated annealing and direct search method for nonlinear unconstrained global optimization. Optimization Methods and Software, to appear.
5. Horst, R., Pardalos,P. M. (Eds.) (1995) Handbook of Global Optimization. Kluwer Academic Publishers, Boston, MA
6. Kelley, C. T. (1999) Detection and remediation of stagnation in the Nelder-Mead algorithm using a sufficient decrease condition. SIAM J. Optim. **10**, 43-55
7. Michalewicz, Z. (1996) Genetic algorithms + Data structures = Evolution programs. Springer, Berlin, Heidelberg, New York
8. Moscato, P. (1999) Memetic algorithms: An introduction. In: Corne, D., Dorigo, M., Glover, F. (Eds.) New Ideas in Optimization. McGraw-Hill, London, UK
9. Nelder, J. A., Mead, R. (1965) A simplex method for function minimization. Comput. J. **7**, 308-313
10. Osman, I. H., Kelly, J. P. (Eds.) (1996) Meta-Heuristics: Theory and Applications. Kluwer Academic Publishers, Boston, MA
11. Yen, J., Liao, J. C., Lee, B., Randolph, D. (1998) A hybrid approach to modeling metabolic systems using a genetic algorithm and simplex method. IEEE Trans. on Syst., Man, and Cybern. B. **28**, 173-191

On Minimax and Maximin Values in Multicriteria Games

Masakazu Higuchi and Tamaki Tanaka

Graduate School of Science and Technology, Niigata University,
Niigata 950-2181, Japan
E-mail: tamaki@m.sc.niigata-u.ac.jp

Abstract. The paper is concerned with a multicriteria game whose payoff takes its values in an ordered vector space. As compare with usual single-criterion games, useful results in such classical games do not always hold as to multicriteria games. For example, minimax and maximin values are coincident under certain conditions in usual single-criterion games, but their values are not always coincident under similar conditions in multicriteria games. Therefore, in this paper, we propose a certain coincidence condition on minimax and maximin values under suitable setting in multicriteria games.

1 Introduction

Game theory started from famous minimax theorem and has been developed widely. In game theory for single-criterion, several kinds of solution concepts and their useful properties have been analyzed and used; see [4] and references cited therein. One of unsolved problems in game theory is whether games with multiple non-comparable criteria have an acceptable theory similar to standard results for single-criterion games. In general, such kind of games is called "multicriteria game", and it has been studied as strategic forms in [1,5,6] and references cited therein. Optimal strategies of a multicriteria game are characterized by a plural number of incomparable optimal solutions, denoted by a set of efficient (or Pareto, or nondominated) points. In single-criterion games, it is well-known that minimax and maximin values are coincident with the saddle value under some conditions, but its analogy for multicriteria games can not be expected in general. Therefore, the aim of this paper is to examine what kind of condition is sufficient for minimax and maximin values to be coincident with each other in multicriteria games.

2 Multicriteria two-person zero-sum game

First, we define a partial ordering for vector-values instead of the total ordering for real-values.

Definition 1. For any two vectors x and y,

$$x \leq_C y \Longleftrightarrow y - x \in C, \quad x <_C y \Longleftrightarrow y - x \in C \backslash \{\theta\}$$
$$x \not\leq_C y \Longleftrightarrow y - x \notin C, \quad x \not<_C y \Longleftrightarrow y - x \notin C \backslash \{\theta\}$$

where C is a solid pointed convex cone, i.e., $\mathrm{int}C \neq \emptyset,\ C \cap (-C) = \{\theta\}$.

Next, we introduce concepts of C-minimal and C-maximal points of a set with respect to the ordering defined by a cone C, i.e., concepts of lower and upper efficient points. Throughout this paper, let Z be an ordered vector space with an ordering $\leq_C$.

Definition 2. $z_0 \in A \subset Z$ is said to be a C-minimal point of A if $A \cap (z_0 - C) = \{z_0\}$, and a C-maximal point of A if $A \cap (z_0 + C) = \{z_0\}$, respectively. We denote the set of such all C-minimal (resp. C-maximal) points of A by $\mathrm{Min}A$ (resp. $\mathrm{Max}A$).

Under these definitions, we consider a game $\Gamma = (X, Y, -f, f)$, where X and Y are nonempty sets, and $f : X \times Y \to Z$. The set X (resp. Y) is the set of strategies of Player 1 (resp. Player 2), and the mapping $-f$ (resp. f) is the payoff function of this player. We call this game "multicriteria two-person zero-sum game", and we can consider the following idea of equilibrium strategies in the same manner as single-criterion games.

Definition 3. In multicriteria two-person zero-sum games, a point (x_0, y_0) is said to be an equilibrium optimal response strategy pair of the game if $f(x, y_0) \not<_C f(x_0, y_0)$ and $f(x_0, y_0) \not<_C f(x_0, y)$, $\forall x \in X,\ y \in Y$.

Above definition is equivalent to the following one.

Definition 4. Let $f : X \times Y \to Z$ be a vector-valued function. A point (x_0, y_0) is said to be a C-saddle point of f with respect to $X \times Y$ if $f(x_0, y_0) \in \mathrm{Max}f(x_0, Y) \cap \mathrm{Min}f(X, y_0)$.

The set of C-saddle values is denoted by $SV(f)$. Let

$$\begin{aligned} D_1 &= \{(x_0, y_0) \in X \times Y \mid f(x_0, y_0) \in \mathrm{Max}f(x_0, Y)\} \quad \text{and} \\ D_2 &= \{(x_0, y_0) \in X \times Y \mid f(x_0, y_0) \in \mathrm{Min}f(X, y_0)\}, \end{aligned}$$

$D_1 \cap D_2$ is the set of all C-saddle points of f, and $f(D_1 \cap D_2) = SV(f)$. By calculating D_1 and D_2, we can easily obtain the set $SV(f)$; see Example 1.

Moreover, by using concepts of C-minimal and C-maximal points, we can define the following subsets of Z as analogues of minimax and maximin values in single-criterion games.

Definition 5. In multicriteria two-person zero-sum games, subsets of Z

$$\mathrm{Minimax}f := \mathrm{Min} \bigcup_{x \in X} \mathrm{Max}f(x, Y) \text{ and } \mathrm{Maximin}f := \mathrm{Max} \bigcup_{y \in Y} \mathrm{Min}f(X, y)$$

are the set of all minimax values for f and the set of all maximin values for f, respectively.

Note that these three sets $SV(f)$, $\mathrm{Minimax}f$ and $\mathrm{Maximin}f$ are not coincident in general.

Example 1. Let Z and C be a 2-dimensional Euclidean space and its positive orthant of Z, respectively. We consider the following game $\Gamma = (X, Y, -f, f)$. X and Y are sets of mixed strategies in another 2-dimensional Euclidean space, i.e., $\text{co}\{(1,0)^t, (0,1)^t\}$. We consider the following matrix type payoff function $f(x,y) = (x^t Ay,\ x^t By)^t$ where

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

This example is given by Corley [1]. Then, we get the set of C-saddle values

$$\begin{aligned} SV(f) &= \{f(x,y) \mid (x,y) \in D\} \\ &= \left\{(u,v)^t \;\middle|\; u = y_1,\ v = -y_1 + 1,\ 0 \le y_1 \le \tfrac{1}{2}\right\} \\ &\quad \cup \left\{(u,v)^t \;\middle|\; \begin{matrix} u = x_1 + y_1 - 2x_1y_1,\ v = -y_1 + x_1y_1 + 1, \\ 0 \le x_1 < \tfrac{1}{2},\ \tfrac{1}{2} < y_1 \le 1 \end{matrix}\right\}, \end{aligned}$$

where $x = (x_1, 1 - x_1)^t,\ y = (y_1, 1 - y_1)^t$ and

$$\begin{aligned} D &:= D_1 \cap D_2 \\ &= \left\{(x,y) \;\middle|\; x_1 = 0,\ 0 \le y_1 \le \tfrac{1}{2}\right\} \cup \left\{(x,y) \;\middle|\; 0 \le x_1 < \tfrac{1}{2},\ \tfrac{1}{2} < y_1 \le 1\right\}. \end{aligned}$$

Minimax$f$ and Maximin$f$ for this example are as follows; (see Fig. 1.)

$$\begin{aligned} \text{Minimax}f &= \{(u,v)^t \mid u = y_1,\ v = 1 - y_1,\ 0 \le y_1 \le 1\}, \\ \text{Maximin}f &= \left\{(u,v)^t \;\middle|\; \begin{matrix} u^2 + 4v^2 - 6u - 8v + 4uv + 5 = 0, \\ \tfrac{1}{2} < u \le 1,\ 0 \le v < \tfrac{3}{4} \end{matrix}\right\}. \end{aligned}$$

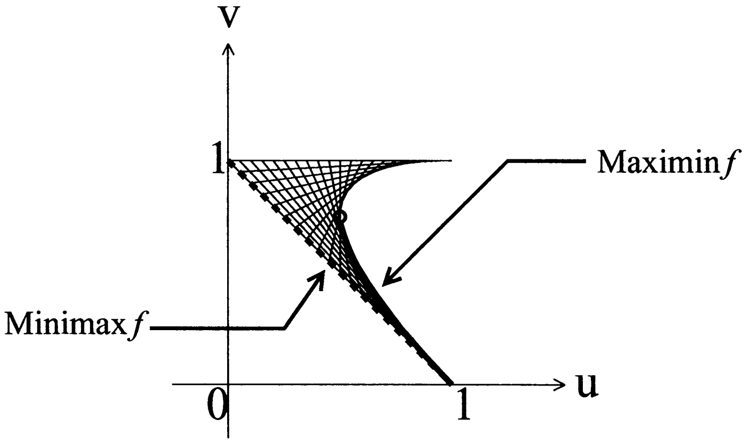


Fig. 1. Minimax$f$ and Maximin$f$ in Example 1.

3 Coincidence condition

In multicriteria two-person zero-sum games, we consider the following setting. Let Z and C be an n-dimensional Euclidean space and its positive orthant of Z, respectively. Strategy sets X and Y are convex hulls generated by $(1,0)^t$ and $(0,1)^t$ in 2-dimensional Euclidean space, i.e., $\text{co}\{(1,0)^t,(0,1)^t\}$. The payoff function f is a bilinear function with respect to x and y.

We introduce a dominance property that is important in the problem of efficient points.

Lemma 1. *(See Lemma 5.2 in [5]) Let Z be an ordered vector space with an ordering defined by a solid pointed convex cone C, and A a subset of Z. If the convex cone C of Z satisfies the condition*

$$\text{cl}C + (C\backslash\{\theta\}) \subset C$$

and if A is nonempty and compact, then $\text{Min}A \neq \emptyset$, $A \subset \text{Min}A + C$ *and* $\text{Max}A \neq \emptyset$, $A \subset \text{Max}A - C$.

As to dominance property, more complex one has been proposed. The first condition in the dominance property is

$$\text{cl}C + (C\backslash L) \subset C$$

where L is the maximal subspace included in C, i.e., $L = C \cap (-C)$, in general; see [2,3]. When C is pointed, this condition is coincident with one of Lemma 1. It is sufficient with this lemma in our setting because Z is the finite-dimensional vector space. By using the dominance property, we can get the following theorem.

Theorem 1. *We assume that $SV(f) \neq \emptyset$ and* $\text{Minimax}f$, $\text{Maximin}f \subset SV(f)$. $\text{Minimax}f = \text{Maximin}f$ *if one of the following statements holds:*

(i) $\forall x \in X,\ d_x \in C \cup (-C)$
(ii) $\forall y \in Y,\ d_y \in C \cup (-C)$

where $d_x = f(x,(1,0)^t) - f(x,(0,1)^t)$ and $d_y = f((1,0)^t,y) - f((0,1)^t,y)$, which are called "direction vector".

Proof. We assume that $d_x \in C\cup(-C)$ for any $x \in X$. For any $z \in \text{Minimax}f$, there exist $x_0 \in X$ and $y_0 \in Y$ such that $z = f(x_0,y_0)$ and

$$z' \not<_C z \text{ and } z \not<_C f(x_0,y), \quad \forall z' \in \text{Max}f(x,Y),\ x \in X,\ y \in Y.$$

Therefore, we have $z \in \text{Max}f(x_0,Y)$. Since we assume that the set of minimax values is a subset of $SV(f)$,

$$z = f(x_0,y_0) \in \text{Max}f(x_0,Y) \cap \text{Min}f(X,y_0),$$

i.e., $f(x, y_0) \not<_C z$, $\forall x \in X$. Since $d_{x_0} \in C \cup (-C)$, we obtain $\text{Max}f(x_0, Y) = \{z\}$. Moreover, C satisfies the condition in Lemma 1 because C is a closed set, and $f(x, Y)$ is a bounded closed set for each $x \in X$, and then it is a compact set. Hence, $f(x_0, Y) \subset z - C$ by Lemma 1. Here, for given $y \in Y$, let $z_{Min(y)}$ be an element of $\text{Min}f(X, y)$. We suppose that $z_{Min(y)} \in z + C \backslash \{\theta\}$, then

$$f(x_0, y) \leq_C z = f(x_0, y_0) \text{ and } z = f(x_0, y_0) <_C z_{Min(y)}.$$

Hence, we obtain $f(x_0, y) <_C z_{Min(y)}$. This is contradictory to $z_{Min(y)} \in \text{Min}f(X, y)$. Therefore, we have $z_{Min(y)} \notin z + C \backslash \{\theta\}$. Since z is also a saddle value,

$$f(x, y_0) \not<_C z \text{ and } z \not<_C z_{Min(y)}, \ \forall z_{Min(y)} \in \text{Min}f(X, y), \ x \in X, \ y \in Y.$$

So, we obtain $z \in \text{Maximin}f$ and hence $\text{Minimax}f \subset \text{Maximin}f$.

On the other hand, for any $z \in \text{Maximin}f$, there exist x_0 and y_0 such that $z = f(x_0, y_0)$ and

$$f(x, y_0) \not<_C z \text{ and } z \not<_C z', \quad \forall z' \in \text{Min}f(X, y), \ x \in X, \ y \in Y.$$

Therefore, we have $z \in \text{Min}f(X, y_0)$. Since we assume that the set of maximin values is a subset of $SV(f)$, we have

$$z = f(x_0, y_0) \in \text{Max}f(x_0, Y) \cap \text{Min}f(X, y_0),$$

i.e., $z \not<_C f(x_0, y)$, $\forall y \in Y$. Here, for given $x \in X$, let $z_{Max(x)}$ be an element of $\text{Max}f(x, Y)$. Then, from $d_x \in C \cup (-C)$, we obtain $\text{Max}f(x, Y) = \{z_{Max(x)}\}$. Moreover, $f(x, Y)$ is a compact set for each $x \in X$ so $f(x, Y) \subset z_{Max(x)} - C$ by Lemma 1. We suppose that $z_{Max(x)} \in z - C \backslash \{\theta\}$, then

$$f(x, y_0) \leq_C z_{Max(x)} \text{ and } z_{Max(x)} <_C z = f(x_0, y_0).$$

Hence, we obtain $f(x, y_0) <_C z = f(x_0, y_0)$. This is contradictory to $z \in \text{Min}f(X, y_0)$. Therefore, we have $z_{Max(x)} \notin z - C \backslash \{\theta\}$. Since z is also a saddle value,

$$z_{Max(x)} \not<_C z \text{ and } z \not<_C f(x_0, y), \ \forall z_{Max(x)} \in \text{Max}f(x, Y), \ x \in X, \ y \in Y.$$

So, we obtain $z \in \text{Minimax}f$ and hence $\text{Minimax}f \supset \text{Maximin}f$.

Consequently, we obtain

$$\text{Minimax}f = \text{Maximin}f.$$

When we also assume that $d_y \in C \cup (-C)$ for any $y \in Y$, we can prove similarly. This completes the proof. □

Note that Theorem 1 holds for the payoff function $f + a$ as well, where a is a vector in Z.

References

1. H. W. Corley. (1985) *Games with Vector Payoffs.* Journal of Optimization Theory and Applications. **47**, 491–498
2. A. S. Karwat. (1986) *On Existence of Cone-Maximal Points in Real Topological Linear Spaces.* Israel Journal of Mathematics. **54**, 33–41
3. D. T. Luc. (1989) *An Existence Theorem in Vector Optimization.* Mathematics of Operations Research. **14**, 693–699
4. L. A. Petrosjan, N. A. Zenkevich (1996) *Game Theory.* World Scientific, Singapore
5. T. Tanaka. (1994) *Generalized Quasiconvexities, Cone Saddle Points, and Minimax Theorem for Vector-Valued Functions.* Journal of Optimization Theory and Applications. **81**, 355–377
6. T. Tanaka. (1994) *Minimax Theorems of Vector-Valued Functions.* Nonlinear analysis and mathematical economics (in Japanese), Kyoto. **861**, 180–196
7. T. Tanaka. (2000) *Vector-Valued Minimax Theorems in Multicriteria Games.* New Frontiers of Decision Making for the Information Technology Era, World Scientific, Singapore. 75–99
8. T. Tanaka, M. Higuchi. (2000) *Classification of Matrix Types for Multicriteria Two-Person Zero-Sum Matrix Games.* Control Applications of Optimization 2000, Pergamon, New York. **2**, 659–668

Backtrack Beam Search for Multiobjective Scheduling Problem

Naoya Honda

Osaka University, 2-1 Yamada-Oka, Suita, Osaka 565-0871 Japan

Abstract. This paper proposes a new approximation method for the multi-objective optimization problem. Proposal method is based on the beam search using the tree structure which is one of the approximate algorithms for combinatorial optimization. In this algorithm, not only finding a solution by search, but backtracking is also executed further, and quality of the solution is improved by using analyzed result of tentative solution. We apply this method to the multi-objective flow-shop problem in which it minimizes maximum completion time, total setup cost and etc. in order to demonstrate the effectiveness of this method by a numerical experiment.

Keywords. Multiobjective Scheduling, Flow-shop, Beam search, Approximation Algorithm, Non-dominated Solution

1. Introduction

Most scheduling problems considered so far treat single objective function only. Scheduling problems of real productions have not a single but multiple criteria. Due to this reason, there is a need to study multiple objective problems.

Flowshop scheduling problems are one of the most well known problems in the scheduling problems. The makespan minimization is often employed as a criterion of flowshop scheduling problem. We treat the multiobjective flowshop problem to minimize two objective, average flow time, and total setup cost, etc. at the same time. As we know, there may not be a schedule that optimizes both criteria at the same time. Thus, we seek non-dominated schedules thereafter.

In the multi-objective problem, it is easy to search only for the solution biased to a certain weight, but it is difficult to search for various solutions not biased. We propose the algorithm which searches for all directions and the solution by using various weights. We also apply our algorithm to the multiobjective flowshop scheduling problem.

This paper is organized as follows. Section 2 defines the non-dominated solution and formulates the problem discussed in this paper. In Section 3, we review the beam search which based on our method and propose our backtrack beam search method. Section 4 applies our method to scheduling problem. The computational experiments and results are given in Section 5. Finally, conclusion and further research problems are discussed in Section 6.

2. Problem Formulation

The definition and the assumption of the scheduling problem (permutation flow shop model) used by our research are shown as follows.

- There are m machines ($M_1, M_2, \dots, M_m$) and n jobs ($J_1, J_2, \dots, J_n$) to be processed on these processors.
- The processing time p_{ij} of job J_i on machine M_j is given.
- For each machine, two jobs or more can not be processed at the same time.
- Each job is processed in the same order of machine, i.e. $M_1, M_2, \dots, M_m$.
- Preemption is not allowed.
- The completion time of job J_i on machine M_j is denoted by c_{ij}, makespan is denoted by $c_i = \max_j c_{ij} = c_{im}$.
- The average flow time is average completion time of all jobs.
- The setup cost is required when changing the job on first machine. The setup cost is not related to time, and does not influence at completion time. The total setup cost is sum of costs between each job.
- The objective function is minimization of makespan, minimization of average flow time, and minimization of total setup cost.
- The purpose of this problem is to search *non-dominated* schedule.
- We define schedule vector v^{π} as a vector consisting two elements, i.e, f_1^{π} and f_2^{π} in some feasible schedule π. That is, $v^{\pi} = (f_1^{\pi}, f_2^{\pi})$.
- For two schedule vectors $v^{\pi_1} = (f_1^{\pi_1}, f_2^{\pi_1})$ and $v^{\pi_2} = (f_1^{\pi_2}, f_2^{\pi_2})$, we say v^{π_1} dominates v^{π_2} when $f_1^{\pi_1} \le f_2^{\pi_1}$, $f_1^{\pi_2} \le f_2^{\pi_2}$ and $v^{\pi_1} \ne v^{\pi_2}$.
- A feasible Schedule π is called to be *non-dominated* if and only if there exists no feasible schedule π' that dominates π.

3. Search Method

3.1. Back-track beam search method

Beam search is a heuristic technique for solving optimization problems. It was adapted from the branch and bound method and was developed in the AI community in the mid 1970's. Lowerre (1976) was the very first person to use this search technique for a speech recognition problem. This technique uses heuristic method to estimate k best promising nodes where k is beam width and hold these k nodes and permanently pruning the rest. The running time of this method is polynomial in the size of the problem compared to exponential in the size of the problems for the branch and bound method.

At each level (except for level 0), k promising nodes are selected for further branching and pruning the rest. The promising nodes are selected based on some evaluation function or criteria and normally related to objective function.

In the beam search, if the accuracy of the evaluation value is not good, the node which contains a good solution might be pruned off. The node pruned off is not searched again. If the beam search is performed without pruned off a node so as to obtain a good solution, it is necessary to expand beam width greatly and search for a lot of nodes at the same time. Therefore, search for solution requires very long calculation time.

In the backtrack beam search, the beam search for the past is performed first of all, evaluation value of the node is calculated and pruned node is stored. Backtrack to the promising node among stored nodes by using the past result so far. The beam search is performed again from the backtracked node. While searching, all nodes pruned are preserved. The lower bound of each node is compared with the tentative solution whether the non-dominated solution exists or not, and the bounded operation is performed to the node without the possibility where non-dominated solution exists.

If backtracking keeps being performed without limitation, this method behaves as the branch and bound method, and a strict solution is obtained. In order to obtain a strict solution, long calculation time is required in a large-scale problem, therefore the searching is stopped at a certain time.

In the multiobjective optimization problem, it is difficult to find various non-dominated solution by searching only at once in the beam search for the past. However, by using backtracking, various non-dominated solution can be found because it is possible to search many times while changing weight. Therefore, our proposal method may be effective in a multiobjective problem.

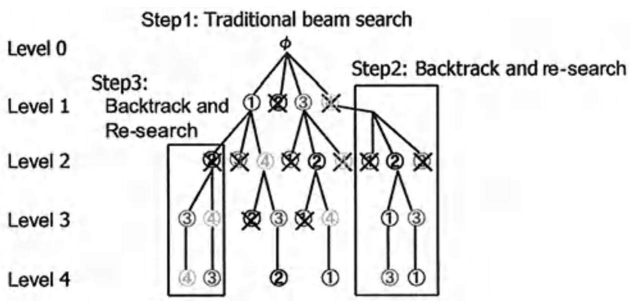


Fig. 1. Search example of backtrack beam search method.

3.2. Lower bound

In order to calculate the evaluation value of the node, lower bound is used. As the lower bound of the makespan, we adopt that of J.Carlier, 1995. As the lower bound of average flow time, we adopt that of S.P.Bansal, 1977. The total setup cost minimization is equivalent to TSP. The lower bound of TSP is given, by the result of J.D.C.Little, 1963.

3.3. Evaluation value

The evaluation value of the node of the beam search is calculated as follows.

$$F = \frac{w_1}{w_1 + w_2} LBf_1 + \frac{w_2}{w_1 + w_2} LBf_2$$

Where weight vector $w = (w_1, w_2)$, the lower bound of objective function 1 is assumed to be LBf_1, the lower bound of objective function 2 is assumed to be LBf_2.

3.4. Algorithms

In a numeric experiment, search stops when the specified calculation time, the repetition search frequency, and the searched number of nodes are reached. Set of solutions at that time is output as an approximation solution.

[Main Procedure]

1. Let n be the number of jobs. Let m be the number of machines. Let r be the number of repetition times at initial phase. Let W be the beam width at initial phase. Let $w = (w_1, w_2)$ be the weight vector at initial phase. $N\{0\}$ is generated as the initial node, and added to the *UnSearch* list. Let L be the un-search node level, and set to $L = 0$.
2. W nodes in level L with good evaluation value are taken out from among the list of *UnSearch*, and those nodes are moved to *SearchList*.
3. *SearchList* is transferred to the function [Beam Search]
4. If non-dominated solution set has been updated, the bounded operation is done to the list of *UnSearch*.
5. If r = 0, go to Step8, otherwise go to Step6.
6. Set $r = r - 1$. *UnSearch* is sorted.
7. The node of the best evaluation value is taken out from among the list of *UnSearch*, the node is moved to *SearchList*, and L is updated to the same level as the node. Go to Step3.
8. Present phase end. It progresses to the next phase; the number of repetition times r, beam width W, and the weight vector w are updated.
9. *UnSearch* is sorted. L is updated to the same level as the highest node. Go to Step2.
10. End of main procedure.

[Beam Search]
1. Let node list *ParentNode* be the received node list. Set $L = L + 1$.
2. Some node is taken out from among the list of *UnSearch*, the node is named N.
3. N is transferred to the function [Expansion of node]. The result is stored in the node list *ChildNode*.
4. If the number of node of *ParentNode* is $|ParentNode| > 0$, go to Step2. Otherwise go to Step5.
5. If the number of node of *ChildNode* is $|ChildNode| = 0$, go to Step11 without updating the non-dominated solution. If $|ChildNode| > 0$, go to Step6.
6. If the searching level $L = n$, go to Step10. Otherwise go to Step7.
7. If $|ChildNode| > W$, node list of *ChildNode* is sorted.
8. W nodes with good evaluation value in *ChildNode* are left, and the remainder is moved to *UnSearch*.
9. All nodes in *ChildNode* are moved to *ParentNode*. Go to Step1.
10. Tentative non-dominated solution is updated by using *ChildNode*.

[Expansion of node]
1. The received node list is named N.
2. The child node is generated with determine the partial job order from the set of un-scheduled job.
3. All generated child nodes are compared with non-dominated solutions. Inferior child node to non-dominated solution is deleted.
4. A remaining node is output as an expanded result set.

4. Numerical Experiments

We apply the proposed algorithm to permutation flow shop problem with bi-criteria, and confirm the effectiveness of our algorithm by numerical experiment. The algorithm has been implemented using C language on an IBM Compatible-PC with 512MB memory and Celeron 533MHz.

The processing time of each job on each machine and setup cost between jobs are randomly generated integer value among interval [0, 50].

Fig.2 shows the non-dominated schedules generated after each phases for the 15 jobs 15 machines problem. This is the result of the transition of the solution among phases. The objective functions are makespan and average flow time. The beam width is 10. The number of backtrackings of each phase is 20. In phase 1, weight vector is (1, 0), this means the priority search of objective function 1. It changes to the next phase after backtracking is perfomed 20 times. Next, the weight vector is (0, 1) in phase 2 to search for objective function 2 by priority. The weight vector from phase 3 to 5 are (2, 1), (1, 2), (1, 1) to search for the balanced solutions. The result shows the effectiveness of the repeated search changing weight.

Fig.3 shows the comparison result of Beam search (BS) and Backtrack beam search (BTBS). The objective functions are makespan and total setup cost. The number of jobs is 20. The number of machine is 20. We show 6 kinds of BS with different weight and one BTBS. Calculation time is the nearly the same respectively. At BS, a beam width is 800 to 1000. In BTBS, a beam width is changed between 5 and 150. Although the solution has been partially biased in BS according to the given weight, the balanced solution is obtained in BTBS.

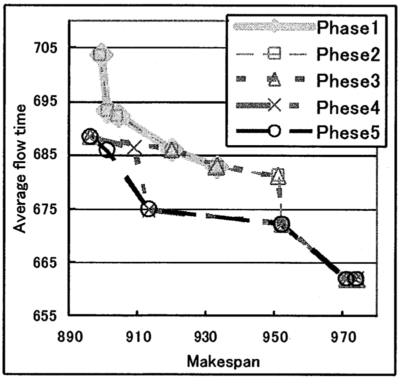


Fig. 2. Non-dominated schedules after each phase

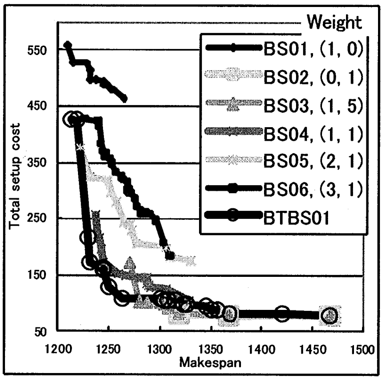


Fig. 3. Beam search vs. Backtrack beam search

5. Conclusion

In this paper, we have introduced an approximation algorithm based on beam search method for multi-objective optimization problems. We have applyed our method for multi-objective flowshop problem which minimize makespan, average flow time and total setup cost. In the proposal algorithm, various solutions can be obtained by the repeated search with changing weight. As shown by the numerical examples, our proposal algorithm can obtain more balanced solution than the beam search method. In our method, it is necessary to give parameters of beam-width, weight, and the number of backtracking, etc. before experiments. We will try the automation of these parameters in the future study.

References

[1] J.Carlier and I.Rebai: Two branch and bound algorithms for the permutation flow shop problem: European Journal of Operational Research, Vol.90, pp. 238-251. (1996)

[2] S.P.Bansal: Minimizing the Sum of Completion Times of n Jobs over m Machines in a Flowshop - A Branch and Bound Approach: AIIE Transactions, Vol.9, pp.306-311. (1977)

[3] J.D.C.Little, G.G.Murty, D.W.Sweeney, C.Karal, "An algorithm for the traveling salesman problem", Operations Research, 11, (1963), 979-989.

Cones to Aid Decision Making in Multicriteria Programming

Brian J. Hunt and Margaret M. Wiecek*

Clemson University, Clemson SC 29634, USA

Abstract. Theoretical and applied aspects of using convex polyhedral cones in multicriteria programming and decision making are explored. Pointed and non-pointed cones are examined and applications of pointed cones to model decision maker's preferences in bicriteria programming are presented.

1 Introduction

A multicriteria program involves a feasible set of decisions evaluated by means of several real-valued criterion functions. As every feasible decision yields an attainable outcome, it is of interest to identify a subset of the feasible decisions producing the best outcomes. In a vast majority of multicriteria problems reported in the literature, the best outcomes are those that outperform each other with respect to the Pareto concept of optimality. This concept has been generalized by many authors. Yu [6] introduced cones to define the best outcomes, which allows for viewing Pareto optimality as a special case of optimality with respect to a convex cone. This was followed by many theoretical and methodological studies on multicriteria programming with convex cones (e.g., [4], [5]). Applied sciences, however, did not follow on this research direction and did not make use of general convex cones in multicriteria problems. The authors of this paper are not aware of a real-life application of multicriteria programming with convex cones. On the other hand, some scientists have undertaken an effort to describe relative importance of criteria with convex cones in order to more accurately model decision maker's preferences (e.g., [2], [3]).

The goal of this paper is to further explore the applicability of convex cones to multicriteria decision making and bring it closer to prospective users. We first connect results by Yu [7] who assumed that the cone used for choosing the best outcomes is acute, and results by Weidner [5] who introduced a more general condition. We then present two applications of polyhedral cones to bicriteria decision making and demonstrate them on an engineering design problem.

* This research was partially supported by the Automotive Research Center, a U.S. Army Center of Excellence for Automotive Research at the University of Michigan.

2 Problem Formulation

The following notation is used throughout this paper. Let $\boldsymbol{y^1}, \boldsymbol{y^2} \in \mathbb{R}^m$ be two vectors. $\boldsymbol{y^1} \geq \boldsymbol{y^2}$ denotes $y_i^1 \geq y_i^2$ for all $i = 1, 2, ..., m$ but $\boldsymbol{y^1} \neq \boldsymbol{y^2}$. $\boldsymbol{y^1} \geqq \boldsymbol{y^2}$ denotes $y_i^1 \geq y_i^2$ for all $i = 1, 2, ..., m$. Also let $\mathbb{R}^m_{\geqq} = \{\boldsymbol{y} \in \mathbb{R}^m | \boldsymbol{y} \geqq \boldsymbol{0}\}$.

Definition 1. A set $C \subseteq R^m$ is called a *cone* if $\boldsymbol{y} \in C$ implies that $\lambda \boldsymbol{y} \in C$ for all $\lambda \geq 0$. A cone C is called *convex* if $\boldsymbol{y^1} \in C$ and $\boldsymbol{y^2} \in C$ implies that $\boldsymbol{y^1} + \boldsymbol{y^2} \in C$. A cone C is called *pointed* if $\boldsymbol{y} \in C, \boldsymbol{y} \neq \boldsymbol{0}$ implies that $-\boldsymbol{y} \notin C$.

A polyhedral cone, $C \subseteq \mathbb{R}^m$, can be represented in intersection form by $\{\boldsymbol{y} \in \mathbb{R}^m | A\boldsymbol{y} \geqq \boldsymbol{0}\}$ where A is a $p \times m$ matrix. Define the nullspace of A as $N(A) = \{\boldsymbol{y} \in \mathbb{R}^m | A\boldsymbol{y} = \boldsymbol{0}\}$.

Let $X \subseteq \mathbb{R}^n$ be a set of all feasible decision vectors and let $f_i(\boldsymbol{x})$, $i = 1, 2, \ldots, m$, be real-valued functions. Then $\boldsymbol{f}(\boldsymbol{x^*})$ represents the attainable outcome vector $[f_1(\boldsymbol{x^*}), \ldots, f_m(\boldsymbol{x^*})]$ for the feasible decision vector $\boldsymbol{x^*}$. Define the set of all attainable outcome vectors as $Y = \{\boldsymbol{y} \in \mathbb{R}^m | \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{x} \in X\} \subseteq \mathbb{R}^m$. Also, define the multiplication of a set $Y \subseteq \mathbb{R}^m$ by a $p \times m$ matrix A as $A \bullet Y = \{\boldsymbol{z} \in \mathbb{R}^p | \boldsymbol{z} = A\boldsymbol{y}, \boldsymbol{y} \in Y\} \subseteq \mathbb{R}^p$ and the algebraic sum of the sets Y and $-Y$ as $Y - Y = \{\boldsymbol{y} \in \mathbb{R}^m | \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x^1}) - \boldsymbol{f}(\boldsymbol{x^2})$ for $\boldsymbol{x^1}, \boldsymbol{x^2} \in X\}$.

Let $C \subset \mathbb{R}^m$ represent a set of "attractive" directions to the decision maker and be referred to as a *preference cone*. A direction is considered "attractive" if traveling along it results in improvement of, or at least no change in, all criterion values with respect to the decision maker's preferences. Yu [6] models the decision maker's preferences with a so-called *domination cone* that contains all directions considered "unattractive" to the decision maker.

Given the problem related to the triple $(X, \boldsymbol{f}, C)$ and following Yu [6], we define efficient decisions and nondominated outcomes with respect to a preference cone C.

Definition 2. A feasible decision $\boldsymbol{x^*} \in X$ is said to be an *efficient decision* for $(X, \boldsymbol{f}, C)$ if there is no direction $\boldsymbol{d} \in C$, $\boldsymbol{d} \neq \boldsymbol{0}$ and no $\boldsymbol{x^1} \in X$ such that $\boldsymbol{f}(\boldsymbol{x^1}) = \boldsymbol{f}(\boldsymbol{x^*}) + \boldsymbol{d}$. An attainable outcome $\boldsymbol{y^*} = \boldsymbol{f}(\boldsymbol{x^*}) \in Y$ is said to be a *nondominated outcome* for $(X, \boldsymbol{f}, C)$ if there is no direction $\boldsymbol{d} \in C$, $\boldsymbol{d} \neq \boldsymbol{0}$ and no $\boldsymbol{y^1} = \boldsymbol{f}(\boldsymbol{x^1}) \in Y$ such that $\boldsymbol{y^1} = \boldsymbol{y^*} + \boldsymbol{d}$. Let $E(X, \boldsymbol{f}, C)$ and $N(Y, C)$ denote the set of efficient decisions and the set of nondominated outcomes, respectively.

3 Pointed and Non-Pointed Cones in Multicriteria Programming

Weidner [5] generalized the concept of pointedness by introducing the condition $N(A) \cap (Y - Y) = \{\boldsymbol{0}\}$ relating the cone to the set of outcomes. We follow on her results and extend them within the framework proposed by Yu [6].

Proposition 1. *If C is a pointed, convex polyhedral cone represented in intersection form, then $N(A) \cap (Y - Y) = \{\mathbf{0}\}$.*

Proof: C is pointed if and only if $rank(A) = m$ which is equivalent to $N(A) = \{\mathbf{0}\}$. By definition, $\mathbf{0} \in (Y - Y)$ so $N(A) \cap (Y - Y) = \{\mathbf{0}\}$. □

Theorem 1. *(Yu [7]) Let C be an acute, convex polyhedral cone represented in intersection form. Then $\boldsymbol{y} \in N(Y, C)$ if and only if $A\boldsymbol{y} \in N(A \bullet Y, \mathbb{R}^p_{\geqq})$.*

Theorem 2. *(Weidner [5]) Let C be a convex polyhedral cone represented in intersection form. Then*

1. *$E(X, \boldsymbol{f}, C) \subseteq E(X, A\boldsymbol{f}, \mathbb{R}^p_{\geqq})$.*
2. *If $N(A) \cap (Y - Y) = \{\mathbf{0}\}$, then $E(X, \boldsymbol{f}, C) = E(X, A\boldsymbol{f}, \mathbb{R}^p_{\geqq})$.*

Proof:
(1) Let $\boldsymbol{x} \in E(X, \boldsymbol{f}, C)$ and assume that $\boldsymbol{x} \notin E(X, A\boldsymbol{f}, \mathbb{R}^p_{\geqq})$. Then there exists a direction $\boldsymbol{d} \in \mathbb{R}^p_{\geqq}$, $\boldsymbol{d} \neq \mathbf{0}$ and an $\boldsymbol{x}' \in X$ such that $A\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{d} = A\boldsymbol{f}(\boldsymbol{x}')$ which yields $\boldsymbol{d} = A\boldsymbol{f}(\boldsymbol{x}') - A\boldsymbol{f}(\boldsymbol{x}) = A(\boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{f}(\boldsymbol{x}))$. Since $\boldsymbol{d} \geq \mathbf{0}$, we know that $A(\boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{f}(\boldsymbol{x})) \geq \mathbf{0}$ and by the representation of the cone C, it must be that $\boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{d}' \in C$. We know that $\boldsymbol{d}' \neq \mathbf{0}$ because $A\boldsymbol{d}' \neq \mathbf{0}$. So there exists a direction $\boldsymbol{d}' \in C$, $\boldsymbol{d}' \neq \mathbf{0}$ and an $\boldsymbol{x}' \in X$ such that $\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{d}' = \boldsymbol{f}(\boldsymbol{x}')$. This means that $\boldsymbol{x} \notin E(X, \boldsymbol{f}, C)$ which is a contradiction. Therefore, $\boldsymbol{x} \in E(X, A\boldsymbol{f}, \mathbb{R}^p_{\geqq})$. □
(2) We show that $E(X, A\boldsymbol{f}, \mathbb{R}^p_{\geqq}) \subseteq E(X, \boldsymbol{f}, C)$. Let $\boldsymbol{x} \in E(X, A\boldsymbol{f}, \mathbb{R}^p_{\geqq})$ and assume that $\boldsymbol{x} \notin E(X, \boldsymbol{f}, C)$. Then there exists a direction $\boldsymbol{d} \in C$, $\boldsymbol{d} \neq \mathbf{0}$ and an $\boldsymbol{x}' \in X$ such that $\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{d} = \boldsymbol{f}(\boldsymbol{x}')$ and also $A\boldsymbol{f}(\boldsymbol{x}) + A\boldsymbol{d} = A\boldsymbol{f}(\boldsymbol{x}')$. Since $\boldsymbol{d} = \boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{f}(\boldsymbol{x})$ then $\boldsymbol{d} \in (Y - Y)$ and because $N(A) \cap (Y - Y) = \{\mathbf{0}\}$ it must be that $\boldsymbol{d} \notin N(A)$ and therefore $A\boldsymbol{d} \neq \mathbf{0}$. However, since $\boldsymbol{d} \in C$, $A\boldsymbol{d} \geqq 0$ or equivalently $A\boldsymbol{d} \in \mathbb{R}^p_{\geqq}$. Let $\boldsymbol{d}' = A\boldsymbol{d}$. Now we have a direction $\boldsymbol{d}' \in \mathbb{R}^p_{\geqq}$, $\boldsymbol{d}' \neq \mathbf{0}$ and an $\boldsymbol{x}' \in X$ such that $A\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{d}' = A\boldsymbol{f}(\boldsymbol{x}')$. This implies that $\boldsymbol{x} \notin E(X, A\boldsymbol{f}, \mathbb{R}^p_{\geqq})$ which is a contradiction. Therefore, $\boldsymbol{x} \in E(X, \boldsymbol{f}, C)$. □

Theorem 3. *Let C be a convex polyhedral cone represented in intersection form. If $N(A) \cap (Y - Y) = \{\mathbf{0}\}$, then $\boldsymbol{y} \in N(Y, C)$ if and only if $A\boldsymbol{y} \in A \bullet N(Y, C)$.*

Proof: ($\Rightarrow$) Obvious. ($\Leftarrow$) Let $A\boldsymbol{y} \in A \bullet N(Y, C)$ and assume that $\boldsymbol{y} \notin N(Y, C)$. Then there exists a $\boldsymbol{y}' \in N(Y, C), \boldsymbol{y}' \neq \boldsymbol{y}$ such that $A\boldsymbol{y} = A\boldsymbol{y}'$ or equivalently $A(\boldsymbol{y} - \boldsymbol{y}') = \mathbf{0}$. This implies that $\boldsymbol{y} - \boldsymbol{y}' \in N(A)$ and because $N(A) \cap (Y - Y) = \{\mathbf{0}\}$, it must be that $\boldsymbol{y} - \boldsymbol{y}' \notin (Y - Y)$. However, $\boldsymbol{y} - \boldsymbol{y}' \in (Y - Y)$ by definition which is a contradiction. Therefore, $\boldsymbol{y} \in N(Y, C)$. □

Theorem 4. *Let C be a convex polyhedral cone represented in intersection form. Then*

1. $A \bullet N(Y, C) \subseteq N(A \bullet Y, \mathbb{R}^p_{\geqq})$.
2. If $N(A) \cap (Y - Y) = \{\mathbf{0}\}$, then $A \bullet N(Y, C) = N(A \bullet Y, \mathbb{R}^p_{\geqq})$.

Proof:
(1) Let $\boldsymbol{z} \in A \bullet N(Y, C)$. Then there exists a $\boldsymbol{y} \in N(Y, C)$ such that $\boldsymbol{z} = A\boldsymbol{y}$. Assume that $\boldsymbol{z} \notin N(A \bullet Y, \mathbb{R}^p_{\geqq})$. Then there exists a direction $\boldsymbol{d} \in \mathbb{R}^p_{\geqq}$, $\boldsymbol{d} \neq \mathbf{0}$ and a $\boldsymbol{z}' \in A \bullet Y$ such that $\boldsymbol{z} + \boldsymbol{d} = \boldsymbol{z}'$. Since $\boldsymbol{z}' \in A \bullet Y$, there exists a $\boldsymbol{y}' \in Y$ such that $\boldsymbol{z}' = A\boldsymbol{y}'$ and $\boldsymbol{d} = \boldsymbol{z}' - \boldsymbol{z} = A\boldsymbol{y}' - A\boldsymbol{y} = A(\boldsymbol{y}' - \boldsymbol{y})$. Since $\boldsymbol{d} \geq \mathbf{0}$, we know that $A(\boldsymbol{y}' - \boldsymbol{y}) \geq \mathbf{0}$ and by the representation of the cone C, it must be that $\boldsymbol{y}' - \boldsymbol{y} = \boldsymbol{d}' \in C$. We know that $\boldsymbol{d}' \neq \mathbf{0}$ because $A\boldsymbol{d}' \neq \mathbf{0}$. So there exists a vector $\boldsymbol{d}' \in C$, $\boldsymbol{d}' \neq \mathbf{0}$ and a $\boldsymbol{y}' \in Y$ such that $\boldsymbol{y} + \boldsymbol{d}' = \boldsymbol{y}'$. This implies $\boldsymbol{y} \notin N(Y, C)$ which is a contradiction. Therefore, $\boldsymbol{z} \in N(A \bullet Y, \mathbb{R}^p_{\geqq})$. □
(2) We show that $N(A \bullet Y, \mathbb{R}^p_{\geqq}) \subseteq A \bullet N(Y, C)$. Let $\boldsymbol{z} \in N(A \bullet Y, \mathbb{R}^p_{\geqq})$. Then there exists a $\boldsymbol{y} \in Y$ such that $\boldsymbol{z} = A\boldsymbol{y}$. Assume that $\boldsymbol{z} \notin A \bullet N(Y, C)$. Then it must be that $\boldsymbol{y} \notin N(Y, C)$ which implies that there exists a $\boldsymbol{d} \in C, \boldsymbol{d} \neq \mathbf{0}$ and a $\boldsymbol{y}' \in Y$ such that $\boldsymbol{y} + \boldsymbol{d} = \boldsymbol{y}'$ and $A\boldsymbol{y} + A\boldsymbol{d} = A\boldsymbol{y}'$. Since $\boldsymbol{d} = \boldsymbol{y}' - \boldsymbol{y}$ then $\boldsymbol{d} \in (Y - Y)$ and because $N(A) \cap (Y - Y) = \{\mathbf{0}\}$ it must be that $\boldsymbol{d} \notin N(A)$ and therefore $A\boldsymbol{d} \neq \mathbf{0}$. However, since $\boldsymbol{d} \in C$, $A\boldsymbol{d} \geqq \mathbf{0}$ or equivalently $A\boldsymbol{d} \in \mathbb{R}^p_{\geqq}$. Let $\boldsymbol{d}' = A\boldsymbol{d}$. Now we have a direction $\boldsymbol{d}' \in \mathbb{R}^p_{\geqq}$, $\boldsymbol{d}' \neq \mathbf{0}$ and a $\boldsymbol{z}' = A\boldsymbol{y}' \in A \bullet Y$ such that $A\boldsymbol{y} + \boldsymbol{d}' = A\boldsymbol{y}'$ or equivalently $\boldsymbol{z} + \boldsymbol{d}' = \boldsymbol{z}'$. This implies that $\boldsymbol{z} \notin N(A \bullet Y, \mathbb{R}^p_{\geqq})$ which is a contradiction. Therefore, $\boldsymbol{z} \in A \bullet N(Y, C)$. □

4 Decision Making With Polyhedral Cones

Pareto-based multicriteria programming uses the nonnegative (nonpositive) orthant to model the decision maker's preferences. Clearly a general convex cone may or may not contain the Pareto cone. From a practical point of view, no direction contained in the Pareto cone should ever be eliminated since every vector in this cone represents a direction in which all criteria increase (decrease) or remain unchanged for a maximization (minimization) problem. Therefore, the general cones we are most interested in using are polyhedral cones that contain the Pareto cone since they contain all the directions of the Pareto cone with some additional directions that in special circumstances may also be attractive to the decision maker.

In the context of decision making, the requirement that the preference cone be pointed translates to the property that if a direction $\boldsymbol{d}$ is attractive to the decision maker then the direction $-\boldsymbol{d}$ should not be considered attractive.

Now consider the bicriteria case and the (nonpositive) Pareto cone. Assume, additionally, that one criterion is relatively more important than the other so that the decision maker allows the value of the less important criterion to decay to obtain an improvement in the value of the more important criterion. This relative importance (RI) of the two criteria can be modeled by opening the Pareto cone into the second quadrant as illustrated in Figure 1.

All directions in the second quadrant model the decision maker's preference that criterion f_1 is relatively more important than criterion f_2 so that it is of interest to improve f_1 even in the presence of decaying f_2. Consistently, the bicriteria problem related to the triple ($X, \boldsymbol{f}$ with RI, Pareto cone) can be equivalently reformulated as the bicriteria problem related to the triple $(X, \boldsymbol{f}, C)$, where C is an RI-related cone.

Another way of interpreting the general cone in bicriteria decision making uses tradeoff information associated with every nondominated outcome. This tradeoff information contains two ratios that inform how much one criterion must be allowed to decay to obtain a unit of improvement in the other criterion. In this sense, constraints may be placed upon the tradeoff information of the nondominated outcomes. Assume that the decision maker is willing to a priori discard those nondominated solutions that do not satisfy certain tradeoff constraints (TOC). Since TOC work in the outcome space they can be used to modify the Pareto cone to a general cone. Consistently, the bicriteria problem related to the triple ($X, \boldsymbol{f}$ with TOC, Pareto cone) can be equivalently reformulated as the bicriteria problem related to the triple $(X, \boldsymbol{f}, C)$, where C is a TOC-related cone.

In view of Section 3, the efficient decisions of the bicriteria problems above can be found by solving related Pareto bicriteria problems for which there are many solution methods.

5 Example

Consider the design of a tractor trailer with two design variables, the height of the hitch (x_1) and the wheelbase of the trailer (x_2). Two criteria of interest in evaluating performance of a tractor trailer are lateral load transfer ratio (f_1) and rearward amplification ratio (f_2) which should both be minimized (see [1] for a detailed model).

Now suppose that the decision maker considers f_1 relatively more important than f_2 and is willing to allow up to 2.3 units of decay in f_2 to gain one unit of improvement in f_1. We could also say that the decision maker considers a tradeoff of at most 2.3 units "attractive".

Alternatively, suppose that the decision maker imposes a TOC that every outcome with a tradeoff of at least 2.3 units is to be retained for consideration. The RI-preference and the TOC-preference of the decision maker are modeled by the obtuse cone in Figure 1, which results in the bicriteria problem related to the triple $(X, \boldsymbol{f}, C)$ where $X = \{(x_1, x_2) \mid x_1 \in [51.2, 76.8], x_2 \in [311.04, 466.56]\}$, $\boldsymbol{f} = [f_1(x_1, x_2), f_2(x_1, x_2)]$, $C = \{\boldsymbol{d} \in \mathbb{R}^2 \mid A\boldsymbol{d} \geqq \boldsymbol{0}\}$ and $A = \begin{bmatrix} -1 & 0 \\ -2.3 & -1 \end{bmatrix}$. In this example C is pointed so by Proposition 1 and Theorem 2, $E(X, \boldsymbol{f}, C) = E(X, -A\boldsymbol{f}, -\mathbb{R}^2_{\geqq}) = E(X, A\boldsymbol{f}, \mathbb{R}^2_{\geqq}) = \{(x_1, x_2) | x_1 \in [59.5, 76.8], x_2 = 466.56\}$. We know that $E(X, \boldsymbol{f}, -\mathbb{R}^2_{\geqq}) = E(X, -\boldsymbol{f}, \mathbb{R}^2_{\geqq}) =$

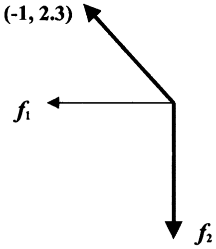


Fig. 1. Pareto cone, RI-related cone and TOC-related cone for the example

$\{(x_1, x_2) \mid x_1 \in [51.2, 76.8], x_2 = 466.56\}$ and we see that the set of efficient decisions is reduced by using the preference cone in Figure 1.

6 Conclusion

In this paper, we investigated the theoretical implications of using non-pointed polyhedral cones in multicriteria programming. We also presented two approaches (relative importance of criteria and tradeoff constraints) to model the decision maker's preferences with pointed polyhedral cones. Problems with polyhedral cones are easily solvable since they can be reformulated as Pareto multicriteria problems.

References

1. Chakarvartula, S., Haque, I., Fadel, G. (to appear 2002) A Monte-Carlo Simulation Approach To Heavy Vehicle Design For Good Dynamic Performance In Multiple Scenarios. International Journal of Heavy Vehicle Systems
2. Noghin, V. D. (1997) Relative Importance of Criteria: A Quantitative Approach. Journal of Multicriteria Decision Analysis **6**:355–363
3. Noghin, V. D., Tolstykh, I. V. (2000) Using Quantitative Information on the Relative Importance of Criteria for Decision Making. Computational Mathematics and Mathematical Physics **40**(11):1529–1536
4. Sawaragi, Y., Nakayama, H., Tanino, T. (1985) Theory of Multiobjective Optimization. Mathematics in Science and Engineering **176**. Academic Press, Inc., Orlando
5. Weidner, P. (1990) Complete Efficiency and Interdependencies Between Objective Functions in Vector Optimization. Methods and Models of Operations Research **34**:91-115
6. Yu, P. L. (1974) Cone Convexity, Cone Extreme Points, and Nondominated Solutions in Decision Problems with Multiobjectives. Journal of Optimization Theory and Applications **14**(3):319–377
7. Yu, P. L. (1985) Multiple Criteria Decision Making: Concepts, Techniques, and Extensions. Plenum Press, New York

Efficiency in Solution Generation based on Extreme Ray Generation Method for Multiple Objective Linear Programming

Masaaki Ida

National Institution for Academic Degrees, 3-29-1 Otsuka, Bunkyo, Tokyo 112-0012, Japan

Abstract. Solution for multiple objective linear programming is a set of all efficient or Pareto optimal solutions. Hence development of useful solution generation method has been desired. We proposed efficient solution generation method based on extreme ray generation method that sequentially generates efficient points and rays by adding inequality constraints of the polyhedral feasible region. In the conventional multiple objective programming researches it is required to solve efficiency test subproblem. On the other hand in our method by investigating the properties of efficiency tests we can improve efficiency test process in solution generation method.

1 Introduction

Definition 1 (Multiple objective linear programming).

$$\begin{array}{ll} \text{maximize} & \boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x} \\ \vdots & \\ \text{maximize} & \boldsymbol{c}_l^{\mathrm{T}}\boldsymbol{x} \end{array} \tag{1}$$

$$\text{subject to} \quad A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0} \tag{2}$$

where $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_l$ are given vectors in $\mathbb{R}^n$, $\boldsymbol{x}$ is a decision variable vector in $\mathbb{R}^n$, A is a given $m \times n$ matrix, and $\boldsymbol{b}$ is a given vector in $\mathbb{R}^m$. We denote $l \times n$ criterion matrix by $C = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_l)^{\mathrm{T}}$ and the polyhedral feasible region by $P = \{\boldsymbol{x} \mid A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}\}$.

Definition 2 (Efficiency). A point $\boldsymbol{x}' \in P$ is said to be *efficient* or *Pareto optimal*, if there does not exists another $\boldsymbol{x} \in P$ such that $C\boldsymbol{x} \geq C\boldsymbol{x}'$, where $\geq$ means at least one strict inequality.

In the past three decades various simplex-like algorithms that sequentially enumerate *efficient extreme points* and *efficient extreme rays* have been developed [1]. On the other hand a new direction of computational techniques in efficient solution generation method based on *extreme ray generation method* or *non-pivoting method* has been developed [2]. The purpose of this paper is to make progress in efficiency tests in objective space to improve the solution generation method.

2 Cone representation and efficiency test

We denote a *convex hull* of a set $S = \{x^1, \ldots, x^u\}$ and a *cone* generated by S as follows:

$$\text{conv } S = \{\sum_{i=1}^{u} \lambda_i x^i \mid \sum_{i=1}^{u} \lambda_i = 1,\ \lambda_i \geqq 0 \text{ for } i = 1, \ldots, u\} \tag{3}$$

$$\text{cone } S = \{\sum_{i=1}^{u} \mu_i x^i \mid \mu_i \geqq 0 \text{ for } i = 1, \ldots, u\} \tag{4}$$

In the polyhedral theory the following proposition for polyhedral set P is stated [3].

Proposition 1. *Any pointed polyhedral set P has a unique minimal representation as*

$$P = \text{conv } U + \text{cone } V \tag{5}$$

where $U = \{u^1, \ldots, u^s\}$ is the set of exactly the all vertices, extreme points, of P, and $V = \{v^1, \ldots, v^t\}$ is the set of nonzero representatives for the all extreme rays of the cone of P.

In order to deal with efficiency condition in cone representation we utilize the following Proposition directly deduced from Definition 2:

Proposition 2. *A feasible solution $x' \in P$ is efficient iff*

$$\text{cone}\{C(u^1 - x'), \ldots, C(u^s - x'),\ Cv^1, \ldots, Cv^t\} \cap \mathbb{R}^l_+ = \{0\}. \tag{6}$$

Hereafter we use the symbol (v^i) to denote the ray $\{\lambda v^i \mid \lambda \geqq 0\}$ by nonzero representative v^i. In the polyhedral feasible region we focus on the extreme feasible solutions such that $x' = u'$ or $x' = u' + (v')$.

Concerning to check efficiency condition by Proposition 2 we obtain the following effective efficiency test problems:

Proposition 3 (Efficiency test). *A feasible solution $x' \in P$ is efficient iff the maximum value of the following linear programming problem is grater than zero:*

$$\max \{z_0 \mid z = \{z_0, z_1, \ldots, z_l\} \in F\} \tag{7}$$

$$\begin{aligned} F = \{z \in \mathbb{R}^{l+1} \mid\ & \sum_{k=1}^{l} z_k (C(u^i - x'))_k \leqq 0 \\ & \textit{for } i \in \{1, \ldots, s\} \setminus \{i' \mid C(u^{i'} - x') \leqq 0\}, \textit{ and} \\ & \sum_{k=1}^{l} z_k (Cv^j)_k \leqq 0 \ \ \textit{for } j \in \{1, \ldots, t\} \setminus \{j' \mid Cv^{j'} \leqq 0\} \\ & \textit{and } z_0 \geqq 0,\ z_i \geqq z_0 \ \ \textit{for } i = 1, \ldots, l\ \} \end{aligned} \tag{8}$$

(*Proof*) The event of Proposition 2 is equivalent to that there does not exist $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{s+t}) \geqq \mathbf{0}$ such that $\sum_{i=1}^{s} \mu_i C(u^i - x') + \sum_{j=1}^{t} \mu_{s+j} C v^j \geq \mathbf{0}$. The alternative of this event is that there exists $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_l) \in \mathbb{R}^l$ such that $\sum_{k=1}^{l} \nu_k (C(\boldsymbol{u}^i - \boldsymbol{x}'))_k \leqq 0$ for $i = 1, \ldots, s$, $\sum_{k=1}^{l} \nu_k (C\boldsymbol{v}^j)_k \leqq 0$ for $j = 1, \ldots, t$, and $\nu_j > 0$ for $j = 1, \ldots, l$. If there exists $\boldsymbol{\nu} \geqq \mathbf{0}$, then $\nu_k (C(\boldsymbol{u}^i - \boldsymbol{x}'))_k \leqq 0$ for $\{i' \mid C(\boldsymbol{u}^{i'} - \boldsymbol{x}') \leqq 0\}$, and similar to the case of extreme rays.

In a similar way to Proposition 3 we obtain the following efficiency test.

Proposition 4 (Efficiency test'). *A feasible solution $\boldsymbol{x}' \in P$ is efficient iff the maximum value of the following linear programming problem is grater than zero:*

$$\max \{z_0 \mid \boldsymbol{z} = \{z_0, z_1, \ldots, z_l\} \in F\} \tag{9}$$

$$\begin{aligned} F = \{\boldsymbol{z} \in \mathbb{R}^{l+1} \mid \ & \sum_{k \notin S_1} z_k (C(\boldsymbol{u}^i - \boldsymbol{x}'))_k \leqq 0, \\ & S_1 = \{k \mid (C(\boldsymbol{u}^i - \boldsymbol{x}'))_k > 0, i = 1, \ldots, s\} \\ & \textit{for } i \in \{1, \ldots, s\}, \textit{ and} \\ & \sum_{k \notin S_2} z_k (C\boldsymbol{v}^j)_k \leqq 0, \\ & S_2 = \{k \mid (C\boldsymbol{v}^j)_k > 0, i = 1, \ldots, t\} \\ & \textit{for } j \in \{1, \ldots, t\} \\ & \textit{and } z_0 \geqq 0,\ z_i \geqq z_0 \ \textit{ for } i = 1, \ldots, l \ \} \end{aligned} \tag{10}$$

Now concerning the extreme points of P we can state the following proposition that is effective in efficiency test process.

Proposition 5. *If $E(\cdots E(P, k_1), \cdots, k_m)$ has only one extreme point $\boldsymbol{u}'$ for $k_1, \cdots, k_m \in \{1, \ldots, l\}$, then $\boldsymbol{u}'$ is efficient.*

where $E(S, k)$ is the solution set $E(S, k)$ in a set S for one objective function $\boldsymbol{c}_k \boldsymbol{x}$ $(k \in \{1, \ldots, l\})$, i.e., $E(S, k) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \arg \max_{\boldsymbol{x} \in S} \boldsymbol{c}_k \boldsymbol{x}\}$, i.e., this proposition means that an optimal solution for each objective function is efficient.

3 Efficient solution generation algorithm

We consider the cone K as cone representation of the feasible region P. Let $(\boldsymbol{w}) = ((\boldsymbol{x}, \xi))$ is an extreme ray of K.

$$K = \{\boldsymbol{w} = (\boldsymbol{x}, \xi) \in \mathbb{R}^{n+1} \mid (-A, \boldsymbol{b})\, \boldsymbol{w} = -A\boldsymbol{x} + \boldsymbol{b}\xi \geqq \mathbf{0},\ \boldsymbol{w} \geqq \mathbf{0}\} \tag{11}$$

If $\xi = 0$, i.e., $\boldsymbol{w} = (\boldsymbol{x}, 0)$, then $(\boldsymbol{x})$ $(= (\boldsymbol{v}))$ is an extreme ray of P, and if $\xi > 0$ and we rewrite the ray as $(\boldsymbol{w}) = ((\boldsymbol{x}/\xi, 1))$, then $\boldsymbol{x}/\xi$ $(= \boldsymbol{u})$ is the extreme point of P.

We concern ourselves with finding all the extreme rays of the form $K = \{\boldsymbol{w} \mid D\boldsymbol{w} \geqq \boldsymbol{0}, \boldsymbol{w} \geqq 0\}$, where D is $n_1 \times n_2$. Consider the matrix $\left(\frac{D}{I}\right)$ where I is $n_2 \times n_2$ identity matrix. *Extreme ray generation method* gives a series of transformation of this matrix that generates all the extreme rays [4]. At any stage of the process we denote the old matrix by $Y = \left(\frac{U}{L}\right)$, and the new matrix being generated denoted by $\bar{Y}$

Now we can utilize the properties of efficiency in polyhedral representation given in Section 2 to improve the efficient solution generation method based on extreme ray generation method [2]:

(Step 0) Set $Y = \begin{pmatrix} -A & \boldsymbol{b} \\ I & \end{pmatrix}$.

(Step 1)

(1) If any row of U has all components negative, then $\boldsymbol{w} = \boldsymbol{0}$ is the only solution of K.

(2) If all the elements of U are nonnegative, then the columns of L are the edges of K, i.e., the ray $(L_{\cdot j})$ is an edge of K.

(Step 2)

(1) Choose the row of U, say row r, with at least one negative elements.

(2) (a) Let $\Psi = \{\, j \mid y_{rj} \geqq 0\}$ and $v = |\Psi|$ (i.e., the number of elements of Ψ). Then the first v columns of the new matrix, $\bar{Y}$, are $Y_{\cdot j}$ $(j \in \Psi)$.

(b) If Y has only two columns and $y_{r1}y_{r2} < 0$, adjoin the column $|y_{r2}|Y_{\cdot 1} + |y_{r1}|Y_{\cdot 2}$ to the $\bar{Y}$ matrix. Go to step 5.

(3) Let $S = \{(s,t) \mid y_{rs}y_{rt} < 0, s < t\}$, i.e., the set of all (unordered) pairs of columns of Y whose elements in row r have opposite signs. Let I_0 be the index set of all nonnegative rows of Y. For each $(s,t) \in S$, find all $i \in I_0$ such that $y_{is} = y_{it} = 0$. Call this set $I_1(s,t)$. We now use some of the elements of S to create additional columns for $\bar{Y}$:

(a) If $I_1(s,t) = \emptyset$, then $Y_{\cdot s}$ and $Y_{\cdot t}$ do not contribute another column to the new matrix.

(b) If $I_1 \neq \emptyset$, check to see if there is a u not equal to either s or t, such that $y_{iu} = 0$ for all $i \in I_1(s,t)$. If such a u exists, then $Y_{\cdot s}$ and $Y_{\cdot t}$ do not contribute another column to the new matrix. If no such u exists, then choose α_1, $\alpha_2 > 0$ to satisfy $\alpha_1 y_{rs} + \alpha_2 y_{rt} = 0$. (one such choice is $\alpha_1 = |y_{rt}|$, $\alpha_2 = |y_{rs}|$.) Adjoin the column $\alpha_1 Y_{\cdot s} + \alpha_2 Y_{\cdot t}$ to the new matrix.

(4) When all pairs in S have been examined, and the additional columns (if any) have been added, we say that row r has been *processed.* Now let Y denote the matrix $\bar{Y}$ produced in *processing* row r.

(Step 3) For each extreme rays $L_{\cdot j}$ of Y at this stage, by noting that $L_{n+1,j} = 0$ means that the column v_j of $L_{\cdot j} = (v_i, 0)$ is the extreme ray of P and

$L_{n+1,j} = 1$ means that the column u_j of $L_{\cdot j} = (u_i, 1)$ is the extreme point of P, discriminate extreme rays $\{v^1, \ldots, v^q\}$ and extreme points $\{u^1, \ldots, u^p\}$.

(Step 4) Check efficiency for $L_{\cdot j}$ coresponding to each extreme ray or point. After checking efficiency for all extreme rays and points of L, go to (Step 1).

Finally we can obtain all efficient extreme points and rays of P.

In each iteration in the algorithm obviously we can state about efficiency as follows:

Proposition 6. *If the pair of extreme rays* $((x^1, \xi^1))$ *and* $((x^2, \xi^2))$ *of* K *such that* x^1 *and* x^2 *are efficient and one of them is eliminated in the row process step, then feasible solution* x' *of the newly generated ray* $((x', \xi'))$ *from them is efficient.*

4 Numerical example

$$C = \begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} -1 & 3 \\ 1 & 3 \\ 4 & 3 \\ 3 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 21 \\ 27 \\ 45 \\ 30 \end{pmatrix}$$

The initial matrix for this problem is denoted by Y^1. The objective values for each extreme ray are represented by corresponding column of the objective value matrix Z^1. Let E and N denote *efficient* and *not efficient*.

$$Y^1 = \begin{bmatrix} 1 & -3 & 21 \\ -1 & -3 & 27 \\ -4 & -3 & 45 \\ -3 & -1 & 30 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \Rightarrow \quad Y^2 = \begin{bmatrix} 0 & 0 & 1 & 21 \\ -6 & 6 & -1 & 27 \\ -15 & 24 & -4 & 45 \\ -10 & 23 & -3 & 30 \\ 3 & 0 & 1 & 0 \\ 1 & 7 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\begin{matrix} N & N & N \end{matrix} \qquad\qquad \begin{matrix} E & E & N & N \end{matrix}$$

$$Z^1 = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad Z^2 = \begin{pmatrix} -1 & 14 & -1 & 0 \\ 7 & 7 & 2 & 0 \end{pmatrix}$$

(1) All extreme rays and extreme point of the initial matrices are *not efficient* by Proposition 2 since there exists positive ray in objective space.

(2) Concerning the second matrices (Y^2, Z^2) we examine the extreme point $(0, 7)^T$ by Proposition 3 and 4 as follows:

$$(z_1, z_2) \begin{pmatrix} -1 & 0 & -1 & -14 \\ 7 & 0 & 2 & -7 \end{pmatrix} \leq \mathbf{0} \Rightarrow (z_1, z_2) \begin{pmatrix} -1 & -1 \\ 7 & 2 \end{pmatrix} \leq \mathbf{0} \Rightarrow (-z_1, -z_2) \leq$$

$\mathbf{0}$, then there exists $z \geq \mathbf{0}$, therefore, the extreme point $(0, 7)$ is efficient.

Next we examine the extreme ray $(0, 7)^T + ((3, 1))^T$ by Proposition 3 as follows:

$$(z_1, z_2)\begin{pmatrix} -1 & 1 & -7 & -13 \\ 7 & -7 & 2 & -14 \end{pmatrix} \leqq \mathbf{0} \Rightarrow (z_1, z_2)\begin{pmatrix} -1 & 1 & -1 \\ 7 & -7 & 2 \end{pmatrix} \leqq \mathbf{0},$$

then there exists $\boldsymbol{z} \geqq \mathbf{0}$, therefore, this extreme ray is efficient.

The third and fourth matrices are as follows:

$$\Rightarrow Y^3 = \begin{bmatrix} 0 & 0 & 21 & 31 \\ -7.8 & 6 & 27 & 17 \\ -10.5 & 24 & 45 & 5 \\ 0 & 23 & 30 & 0 \\ 6.9 & 0 & 0 & 10 \\ 9.3 & 7 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \Rightarrow \quad Y^4 = \begin{bmatrix} 0 & 0 & 21 & 21 & 31 \\ -3.6 & 6 & 9 & 27 & 17 \\ 0 & 24 & 0 & 45 & 5 \\ 7 & 23 & 0 & 30 & 0 \\ 4.8 & 0 & 9 & 0 & 10 \\ 8.6 & 7 & 3 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\begin{matrix} E & E & N & N \end{matrix} \qquad \begin{matrix} E & E & E & N & N \end{matrix}$$

$$\Rightarrow Z^3 = \begin{pmatrix} 11.7 & 14 & 0 & -10 \\ 23.1 & 7 & 0 & 20 \end{pmatrix} \Rightarrow Z^4 = \begin{pmatrix} 12.4 & 14 & -3 & 0 & -10 \\ 18.2 & 7 & 21 & 0 & 20 \end{pmatrix}$$

(3) In the third matrices (Y^2, Z^2), the extreme point $(6.9, 9.3)$ is efficient by Proposition 5 and the extreme point $(10, 0)$ is not efficient by Proposition 2.
(4) The newly generated extreme point $(4.8, 8.6)$ is efficient by Proposition 6 and $(9, 3)$ is efficient by Proposition 5.
(5) Finally the extreme points (3, 8) and (6, 7) are efficient by Proposition 6.

$$\Rightarrow \quad Y^5 = \begin{bmatrix} 0 & 0 & 6 & 21 & 21 & 31 \\ 0 & 6 & 0 & 9 & 27 & 17 \\ 9 & 24 & 0 & 0 & 45 & 5 \\ 13 & 23 & 5 & 0 & 30 & 0 \\ 3 & 0 & 6 & 9 & 0 & 10 \\ 8 & 7 & 7 & 3 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad Z^5 = \begin{matrix} E & E & E & E & N & N \\ \end{matrix} \begin{pmatrix} 13 & 14 & 8 & -3 & 0 & -10 \\ 14 & 7 & 19 & 21 & 0 & 20 \end{pmatrix}$$

5 Conclusion

In this paper, we consider the properties of efficiency tests in objective space in a new direction of efficient solution generation method. More considerations about comparison between the proposed method and the conventional pivoting methods are expected.

References

1. Steuer, R. E. (1986) Multiple Criteria Optimization: Theory, Computation, and Application. John Wiley and Sons
2. Ida, M. (1998) Efficient Solution Generation for Multiobjective Linear Programming based on Non-Pivoting Method. Transactions of the Society of Instrument and Control Engineers (Japanese), **34**, 866–868
3. Schrijver, A. (1986) Theory of Liner and Integer Programming. John Wiley and Sons
4. Matheiss, T. H. and Rubin, D. S. (1980) A Survey and Comparison of Methods for Finding All Vertices of Convex Polyhedral Sets. Mathematics of Operations Research, **5**, 167–185

Robust Efficient Basis of Interval Multiple Criteria and Multiple Constraint Level Linear Programming

Masaaki Ida

National Institution for Academic Degrees, 3-29-1 Otsuka, Bunkyo, Tokyo 112-0012, Japan

Abstract. In this paper we deal with interval multiple criteria and multiple constraint level linear programming. We define a robust basis for all possible perturbation of coefficients within intervals in objective functions and constraints that is regarded as secure and conservative solution under uncertainty. According to the conventional multiple objective programming literature, it is required to solve test subproblem for each basis. Therefore, in case of our interval problem excessive computational demand is estimated. In this paper investigating the properties of robust basis by combination of interval extreme points we obtained the result that the robust basis can be identified by working with only a finite subset of possible perturbations of the coefficients.

1 Introduction

Generally it is difficult to determine exactly the coefficients in mathematical programming problems due to various kinds of uncertainties. However, it is sometimes possible to estimate the perturbations of coefficients by intervals or possibilistic distributions. For such decision making situations, interval mathematical programming or fuzzy mathematical programming with uncertain coefficients have been investigated [1],[2],[3],[4]. In the setting of fuzzy multiple objective programming with possibilistic coefficients two kinds of efficient solution sets are defined as fuzzy sets. In the interval case where all possibilistic coefficients degenerate into interval coefficients, important results for two kinds of efficiency tests were obtained [5],[6],[7], i.e., efficient solutions can be identified by finite subsets of the possible perturbations of the coefficients in the interval matrix.

In this paper more general results are obtained in the framework of interval multiple criteria and multiple constraint level linear programming. We define a robust potential basis for all possible perturbation of coefficients within intervals in objective functions and constraints that is regarded as secure and conservative solution under uncertainty. Investigating the theoretical aspects of robust potential basis We discuss that the robust basis can be identified by working with only a finite subset of possible perturbations of the coefficients.

2 Multiple criteria and multiple constraint level linear programming

A multiple criteria and multiple constraint level linear programming is defined as follows (e.g., [8]):

Definition 1 (*P*1).

$$\text{Minimize } C\boldsymbol{x} \tag{1}$$
$$\text{subject to } A\boldsymbol{x} \geqq D,\ \boldsymbol{x} \geqq \mathbf{0} \tag{2}$$

where C is a $p \times n$ matrix, A is a $m \times n$ matrix, and D is a $m \times q$ matrix.

A potential solution for this problem is defined as follows:

Definition 2. We call $\boldsymbol{x}$ a potential solution that is an optimal solution for the following linear programming problem with $\boldsymbol{\mu} \geq \mathbf{0}$ and $\boldsymbol{\nu} \geq \mathbf{0}$:

$$\text{minimize } \boldsymbol{\nu}^{\mathrm{T}} C\boldsymbol{x} \tag{3}$$
$$\text{subject to } A\boldsymbol{x} \geqq D\boldsymbol{\mu},\ \boldsymbol{x} \geqq \mathbf{0},\ \boldsymbol{\nu} \geq \mathbf{0},\ \boldsymbol{\mu} \geq \mathbf{0} \tag{4}$$

where inequality $\geq$ means at least one strict inequality.

Inequality in constraint condition of this problem is rewritten as

$$A\boldsymbol{x} - \boldsymbol{\lambda} = D\boldsymbol{\mu} \tag{5}$$

Similarly inequality in condition of the dual problem for this problem, dual feasibility, can be represented as the following equation.

$$\boldsymbol{\mu} + A^{\mathrm{T}}\boldsymbol{y} = C^{\mathrm{T}}\boldsymbol{\nu} \tag{6}$$

By using the property of linear programming problem we can represent the optimality condition for our problem.

Proposition 1. *$\boldsymbol{x}$ ($\boldsymbol{y}$) is a potential solution for the problem (**P1**) if it satisfies the following conditions:*

$$(A, -I_{m\times m})\boldsymbol{v} = D\boldsymbol{\mu},\ \boldsymbol{v} \geqq \mathbf{0},\ \boldsymbol{\mu} \geq \mathbf{0} \tag{7}$$
$$(I_{n\times n}, A^{\mathrm{T}})\boldsymbol{w} = C^{\mathrm{T}}\boldsymbol{\nu},\ \boldsymbol{w} \geqq \mathbf{0},\ \boldsymbol{\nu} \geq \mathbf{0} \tag{8}$$
$$\boldsymbol{v} \cdot \boldsymbol{w} = \mathbf{0} \tag{9}$$

where $\boldsymbol{v} = (x_1, \cdots, x_n, \lambda_1, \cdots, \lambda_m)$, $\boldsymbol{w} = (\mu_1, \cdots, \mu_n, y_1, \cdots, y_m)$ and I is an identity matrix.

We discuss a potential basis for our problem in basic form. Let B be an m-tuple of integers from $\{1, \ldots, m+n\}$ called basis, and $N = \{1, \ldots, m+n\} \backslash B$. Let $\boldsymbol{v} = (\boldsymbol{v}_B, \boldsymbol{v}_N)$ and $\boldsymbol{w} = (\boldsymbol{w}_N, \boldsymbol{w}_B)$, where $\boldsymbol{v}_B = \{v_i \mid i \in B\}$, $\boldsymbol{v}_N = \{v_i \mid i \in N\}$, $\boldsymbol{w}_N = \{w_i \mid i \in B\}$, $\boldsymbol{w}_B = \{w_i \mid i \in N\}$. Then we represent a potential basis in a basic form.

Proposition 2. *Basis B is a potential basis if the following conditions are satisfied:*

$$A_1 \boldsymbol{v}_B = D\boldsymbol{\mu},\ \ \boldsymbol{\mu} \geq \mathbf{0} \tag{10}$$

$$A_2 \boldsymbol{w}_B = C^{\mathrm{T}}\boldsymbol{\nu},\ \ \boldsymbol{\nu} \geq \mathbf{0} \tag{11}$$

$$\boldsymbol{v}_B \geqq \mathbf{0},\ \ \boldsymbol{w}_B \geqq \mathbf{0},\ \ \boldsymbol{v}_N = \boldsymbol{w}_N = \mathbf{0} \tag{12}$$

where A_1 is a matrix with the column vectors from $(A, -I_{m\times m})$ corresponding to $\boldsymbol{v}_B$, and A_2 is a matrix with the column vectors from $(I_{n\times n}, A^{\mathrm{T}})$ corresponding to $\boldsymbol{w}_B$.

3 Interval coefficient problem

From a practical point of view due to various kinds of uncertainties it is usually difficult to specify the coefficients of the objective functions and constraints. However, there exist some cases where coefficients can be specified by possible ranges represented by intervals.

In this paper regarding the uncertainties represented by intervals, we consider interval multiple criteria and multiple constraint level linear programming problems.

Definition 3.

$$\text{Minimize}\ \ C\boldsymbol{x} \tag{13}$$

$$\text{subject to}\ \ A\boldsymbol{x} \geqq D,\ \ \boldsymbol{x} \geqq \mathbf{0} \tag{14}$$

where C is an element of a set of $p \times n$ criteria matrix with elements $c_{ij} \in [\,\underline{c}_{ij}, \overline{c}_{ij}\,]$ $(i = 1, \ldots, p,\ \ j = 1, \ldots, n)$:

$$C \in \begin{pmatrix} [\underline{c}_{11}, \overline{c}_{11}] & \cdots & [\underline{c}_{1n}, \overline{c}_{1n}] \\ \vdots & & \vdots \\ [\underline{c}_{p1}, \overline{c}_{p1}] & \cdots & [\underline{c}_{pn}, \overline{c}_{pn}] \end{pmatrix}, \tag{15}$$

A is an element of a set of $m \times n$ matrix with elements $a_{ij} \in [\,\underline{a}_{ij}, \overline{a}_{ij}\,]$ $(i = 1, \ldots, m,\ \ j = 1, \ldots, n)$:

$$A \in \begin{pmatrix} [\underline{a}_{11}, \overline{a}_{11}] & \cdots & [\underline{a}_{1n}, \overline{a}_{1n}] \\ \vdots & & \vdots \\ [\underline{a}_{m1}, \overline{a}_{m1}] & \cdots & [\underline{a}_{mn}, \overline{a}_{mn}] \end{pmatrix}, \tag{16}$$

and D is an element of a set of $m \times q$ matrix with elements $d_{ij} \in [\,\underline{d}_{ij}, \overline{d}_{ij}\,]$ $(i = 1, \ldots, m,\ \ j = 1, \ldots, q)$:

$$D \in \begin{pmatrix} [\underline{d}_{11}, \overline{d}_{11}] & \cdots & [\underline{d}_{1q}, \overline{d}_{1q}] \\ \vdots & & \vdots \\ [\underline{d}_{m1}, \overline{d}_{m1}] & \cdots & [\underline{d}_{mq}, \overline{d}_{mq}] \end{pmatrix}. \tag{17}$$

This problem can be regarded as a set of multiple criteria and multiple constraint level linear programming problems each of which has a matrix C, A, and D in the interval matrices respectively.

For this kind of interval coefficient problems, two kinds of solution concepts, i.e., optimistic and pessimistic solutions, have been investigated [1],[5],[6]. In this paper we define a ***robust potential basis*** as pessimistic or secure solution.

Definition 4. We call B a robust potential basis, if it is a potential basis for all $c_{kj} \in [\, \underline{c}_{kj}, \overline{c}_{kj} \,]$, $a_{ij} \in [\, \underline{a}_{ij}, \overline{a}_{ij} \,]$ and $d_{il} \in [\, \underline{d}_{il}, \overline{d}_{il} \,]$.

4 Main results

According to Proposition 2 we can represent a robust potential basis in a basic form.

Proposition 3. *Basis B is a robust potential basis if the following conditions are satisfied for all $c_{kj} \in [\, \underline{c}_{kj}, \overline{c}_{kj} \,]$, $a_{ij} \in [\, \underline{a}_{ij}, \overline{a}_{ij} \,]$ and $d_{il} \in [\, \underline{d}_{il}, \overline{d}_{il} \,]$:*

$$A_1 \boldsymbol{v}_B = D\boldsymbol{\mu}, \ \boldsymbol{\mu} \geq \mathbf{0} \tag{18}$$
$$A_2 \boldsymbol{w}_B = C^{\mathrm{T}}\boldsymbol{\nu}, \ \boldsymbol{\nu} \geq \mathbf{0} \tag{19}$$
$$\boldsymbol{v}_B \geqq \mathbf{0}, \ \boldsymbol{w}_B \geqq \mathbf{0}, \ \boldsymbol{v}_N = \boldsymbol{w}_N = \mathbf{0} \tag{20}$$

Unfortunately the cardinality of this subset, combination of lower and upper bound of intervals is $2^{pn+mn+qn}$.

Now we define the following two matrix sets:

Definition 5 (Matrix set M_1). We denote a subset by M_1 for (A_1, D) having all elements of each row at the upper bound or at the lower bound. Hence, if $(A_1, D) \in M_1$, for $j = 1, \ldots, m$ either $A_{1i\cdot} = \underline{A_1}_{i\cdot}$, $D_{i\cdot} = \underline{D}_{i\cdot}$, or $A_{1i\cdot} = \overline{A_1}_{i\cdot}$, $D_{i\cdot} = \overline{D}_{i\cdot}$. The maximum number of elements of M_1 is 2^m.

Definition 6 (Matrix set M_2). We denote a subset by M_2 for (A_2, C) having all elements of each column at the upper bound or at the lower bound. Hence, if $(A_2, C) \in M_2$, for $i = 1, \ldots, n$ either $A_{2\cdot j} = \underline{A_2}_{\cdot j}$, $C_{\cdot j} = \underline{C}_{\cdot j}$ or $A_{2\cdot j} = \overline{A_2}_{\cdot j}$, $C_{\cdot j} = \overline{C}_{\cdot j}$. The maximum number of elements of M_2 is 2^n.

Then finally we obtain the following Theorem:

Theorem 1. *Basis B is a robust potential basis if the following conditions are satisfied for every $(A_1, D) \in M_1$ and every $(A_2, C) \in M_2$:*

$$A_1 \boldsymbol{v}_B = D\boldsymbol{\mu}, \ \boldsymbol{\mu} \geq \mathbf{0} \tag{21}$$
$$A_2 \boldsymbol{w}_B = C^{\mathrm{T}}\boldsymbol{\nu}, \ \boldsymbol{\nu} \geq \mathbf{0} \tag{22}$$
$$\boldsymbol{v}_B \geqq \mathbf{0}, \ \boldsymbol{w}_B \geqq \mathbf{0}, \ \boldsymbol{v}_N = \boldsymbol{w}_N = \mathbf{0} \tag{23}$$

Proof. By the theorem of alternative (Motzkin's theorem [9]) occurrence of the event of equation (21) and the event of following inequalities (24) is exclusive.

$$A_1{}^T\zeta \geqq \mathbf{0}, \quad D^{\mathrm{T}}\zeta > \mathbf{0} \tag{24}$$

If A_1 with elements of $a_{ij} \in [\, \underline{a}_{ij}, \overline{a}_{ij}\,]$ and D with elements of $d_{il} \in [\, \underline{d}_{il}, \overline{d}_{il}\,]$ satisfy the inequalities (24), then by the definition of M_1 there exist matrices $(A_1^*, D^*) \in M_1$ such that

$$A_1^{*T}\zeta \geqq A_1{}^T\zeta \geqq \mathbf{0}, \quad D^{*\mathrm{T}}\zeta \geqq D^{\mathrm{T}}\zeta > \mathbf{0} \tag{25}$$

Therefore, it is sufficient to consider the matrix set M_1. Similarly we can discuss the equation (22) with the matrix set M_2.

The cardinality of combination of these subsets is 2^{m+n}. This theorem can be regarded as an extension of the past results for the problem with interval coefficients [5],[6].

5 Conclusion

In this paper considering the optimality condition in linear programming problem we investigated the properties of robust basis for multiple criteria and multiple constraint level linear programming with interval coefficients. By means of the obtained Theorem 1, robust potential basis can be identified by working with only a finite subset of possible perturbations of the coefficients.

References

1. Inuiguchi, M. and Sakawa, M. (1996) Possible and Necessary Efficiency in Possibilistic Multiobjective Linear Programming Problems and Possible Efficiency Test, Fuzzy Sets and Systems, **78**, 231–241
2. Inuiguchi, M. and Ramik, J. (2000) Possibilistic Linear Programming: a Brief Review of Fuzzy Mathematical Programming and a Comparison with Stochastic Programming in Portfolio Selection Problem, Fuzzy Sets and Systems, **115**, 3–28
3. Ida, M. (1995) Optimality on Possibilistic Linear Programming with Normal Possibility Distribution Coefficient, Japanese Journal of Fuzzy Theory and Systems, Allerton Press Inc., **7**, 349–360
4. Ida, M. (1999) Possibility Degree and Sensitivity Analysis in Possibilistic Multiobjective Linear Programming Problems, Proc. of the 8th IEEE International Conference on Fuzzy Systems, **1**, 22–27
5. Bitran, G.R. (1980) Linear Multiple Objective Problems with Interval Coefficients, Management Science, **26**, 694–706
6. Ida, M and Katai, O. (1993) Discrimination Methods of Efficient Solutions for Multiobjective Linear Programming Problems with Interval Coefficients, Trans. of the Soc. of Instrument and Control Engineers Japan, **29**, 1247–1249

7. Ida, M., (2001) Mean-variance Portfolio Optimization Model with Uncertain Coefficients, Proc. of FUZZ-IEEE2001
8. Nemhauser, G.L., Kan, A.H.G. Rinnooy, and Todd, M.J. (eds.) (1989) Handbooks in Operations Research and Management Science, 1: Optimization, North-Holland
9. Mangasarian, O. L. (1969) Nonlinear Programming, McGraw-Hill

An Interactive Satisficing Method through the Variance Minimization Model for Fuzzy Random Multiobjective Linear Programming Problems

Hideki Katagiri[1], Masatoshi Sakawa[1], and Shuuji Ohsaki[1]

Hiroshima University, Kagamiyama 1-4-1, Hiagashi-Hiroshima, Hiroshima, 739-8527 Japan

Abstract. In this paper, we consider a multiobjective linear programming problem involving fuzzy random variable coefficients. Introducing a fuzzy goal for each objective function, we focus on a degree of possibility that each objective function satisfies the corresponding fuzzy goal. Since the degree of possibility varies randomly, we formulate the multiobjective problem to minimize the variances of degrees of possibility based on stochastic programming. In order to find a satisficing solution for a decision maker, we propose an interactive satisficing method based on the reference point method.

1 Introduction

In classical mathematical programming, the coefficients of objectives or constraints in problems are assumed to be completely known. However, in real systems, they are rather uncertain than constant. In order to deal with such uncertainty, stochastic programming [1] and fuzzy programming [2,3] were considered. They are useful tools for the decision making under a stochastic environment or a fuzzy environment, respectively.

Most researches in respect to mathematical programming take account of either fuzziness or randomness. However, in practice, decision makers face with the situations where both fuzziness and randomness exist. For instance, in the case where some expert estimates coefficients of objective functions or constraints with uncertainty, they are not always given as random variables or fuzzy sets but as the values including both fuzziness and randomness. Fuzzy random variables [4,5] are one of the mathematical concepts dealing with fuzziness and randomness simultaneously. Recently, several authors considered linear programming problems involving fuzzy random variables [6–8]. In our previous research, we considered multiobjective fuzzy random linear programming problem[9] using the concept of possibility measures[10] and the expectation model, which is maximize the expectation of degrees of possibility that the objective function values satisfy fuzzy goals. In this research, we consider the problem based on the V-model in stochastic programming[4] and propose an interactive satisficing method in order to obtain a satisficing solution for a decision maker.

2 Formulation

In this paper, we consider the following multiobjective linear programming problem:

$$\left.\begin{array}{l}\min \bar{\tilde{\boldsymbol{c}}}_i \boldsymbol{x}, \quad i=1,\dots,k \\ \text{s. t. } A\boldsymbol{x} \le \boldsymbol{b},\ \boldsymbol{x} \ge \boldsymbol{0}\end{array}\right\} \tag{1}$$

where $\boldsymbol{x} = (x_1, \dots, x_n)^t$ is a decision vector and $\bar{\tilde{\boldsymbol{c}}}_i = (\bar{\tilde{c}}_{i1}, \dots, \bar{\tilde{c}}_{in})$ is a coefficient vector. Let A be an $m \times n$ matrix and $\boldsymbol{b}$ an $m \times 1$ vector. Each $\bar{\tilde{c}}_{ij}$ is a fuzzy random variable with the following membership function:

$$\mu_{\bar{\tilde{C}}_{ij}}(t) = \max\left\{0, 1 - \frac{|t - \bar{c}_{ij}|}{\alpha_{ij}}\right\},\ i = 1, \dots, k,\ j = 1, \dots, n \tag{2}$$

where $\bar{c}_{ij}$ denotes a random variable (or a scenario variable) whose realization under the scenario s_i is c_{ijs_i}, and the number of scenarios s_i corresponding to the ith objective function is S_i. Let p_{is_i} be the probability that each scenario s_i occurs. We assume that $\sum_{s_i=1}^{S_i} p_{is_i} = 1$ holds. Each α_{ij} denotes the spread of a fuzzy number. This type of fuzzy random variable is equivalent to a *hybrid number*, which was introduced by Kaufman and Gupta [12].

Since the coefficients of objective functions are the symmetric triangular fuzzy random variables, each objective function also becomes the same type of fuzzy random variable $\bar{\tilde{Y}}_i$ with the following membership function:

$$\mu_{\bar{\tilde{Y}}_i}(y) = \max\left\{0\ ,\ 1 - \frac{\left|y - \sum\limits_{j=1}^{n} \bar{c}_{ij} x_j\right|}{\sum\limits_{j=1}^{n} \alpha_{ij} x_j}\right\},\ i = 1, \dots, k. \tag{3}$$

Considering the imprecision or fuzziness of the decision maker's judgment, for each objective function of problem (1), we introduce the fuzzy goal $\tilde{G}_i$ with the membership function expressed as

$$\mu_{\tilde{G}_i}(y) = \begin{cases} 0, & y > h_i^0 \\ \dfrac{y - h_i^0}{h_i^1 - h_i^0}, & h_i^1 \le y \le h_i^0 \\ 1, & y < h_i^1, \qquad i = 1, \dots, k. \end{cases} \tag{4}$$

Since the membership function $\mu_{\bar{\tilde{Y}}_i}$ is regarded as a possibility distribution, the degree of possibility $\Pi_{\bar{\tilde{Y}}_i}(\tilde{G}_i)$ that the objective function value satisfies the fuzzy goal $\tilde{G}_i$ is

$$\Pi_{\bar{\tilde{Y}}_i}(\tilde{G}_i) = \sup_y \min\left\{\mu_{\bar{\tilde{Y}}_i}(y),\ \mu_{\tilde{G}_i}(y)\right\},\ i = 1, \dots, k. \tag{5}$$

Accordingly, we consider the following multiobjective problem:

$$\left.\begin{array}{l}\max \Pi_{\bar{\tilde{Y}}_i}(\tilde{G}_i), \quad i=1,\dots,k \\ \text{s. t. } A\boldsymbol{x} \le \boldsymbol{b},\ \boldsymbol{x} \ge \boldsymbol{0}\end{array}\right\} \tag{6}$$

In this research, taking account of all scenarios, we set h_i^0 and h_i^1 as the following form:

$$h_i^0 = \max_{s_i} \max_{\boldsymbol{x} \in X} \sum_{j=1}^{n} c_{ijs_i} x_j, \quad i=1,\dots,k,$$

$$h_i^1 = \min_{s_i} \min_{\boldsymbol{x} \in X} \sum_{j=1}^{n} c_{ijs_i} x_j, \quad i=1,\dots,k,$$

where $X \triangleq \{\boldsymbol{x} | A\boldsymbol{x} \le \boldsymbol{b},\ \boldsymbol{x} \ge \boldsymbol{0}\}$. By using (3) and (4), the degree of possibility is represented as follows:

$$\Pi_{\bar{\tilde{Y}}_i}(\tilde{G}_i) = \frac{\sum_{j=1}^{n} \{\alpha_{ij} - \bar{c}_{ij}\} x_j + h_i^0}{\sum_{j=1}^{n} \alpha_{ij} x_j - h_i^1 + h_i^0}, \quad i=1,\dots,k.$$

Since the degree of possibility in problem (6) varies randomly, the problem is regarded as a stochastic programming problem. Katagiri et al.[9] proposed a fuzzy random multiobjective linear programming model, which is to maximize the expected degree of possibility that objective function values satisfy the respective fuzzy goals. This model is useful for decision making under fuzzy stochastic environments; however, in the obtained solution based on this model, there is a possibility that the degree of possibility corresponding to a certain scenario is fairly small because the variance of the degree of possibility is unconsidered. Therefore, in this research, we propose the model to minimize the variances of degrees of possibility. For $i = 1,\dots,k$, the variances of degrees of possibility are calculated as follows:

$$V[\Pi_{\bar{\tilde{Y}}_i}(\tilde{G}_i)] = \frac{1}{\left(\sum_{j=1}^{n} \alpha_{ij} x_j - h_i^1 + h_i^0\right)^2} V\left[\sum_{j=1}^{n} \bar{c}_{ij} x_j\right]$$

Let V_i denote the variance-covariance matrix of $\bar{\boldsymbol{c}}_i$. Then the problem to minimize the variances of degrees of possibility is formulated as

$$\left.\begin{array}{l}\min \dfrac{1}{\left(\sum_{j=1}^{n} \alpha_{ij} x_j - h_i^1 + h_i^0\right)^2} \boldsymbol{x}^T V_i \boldsymbol{x}, \quad i=1,\dots,k \\ \text{s. t. } A\boldsymbol{x} \le \boldsymbol{b},\ \boldsymbol{x} \ge \boldsymbol{0}\end{array}\right\} \tag{7}$$

The variance-covariance matrix is expressed by

$$V_i = \begin{bmatrix} v_{11}^i & v_{12}^i & \cdots & v_{1n}^i \\ v_{21}^i & v_{22}^i & \cdots & v_{2n}^i \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1}^i & v_{n2}^i & \cdots & v_{nn}^i \end{bmatrix}, \quad i = 1, \dots, k$$

where

$$v_{jj}^i = V[\bar{c}_{ij}] = \sum_{s_i=1}^{S_i} p_{is_i} \{c_{ijs_i}\}^2 - \left\{ \sum_{s_i=1}^{S_i} p_{is_i} c_{ijs_i} \right\}^2, \ j = 1, \dots, n$$

$$v_{jl}^i = Cov[\bar{c}_{ij}, \bar{c}_{il}] = E[\bar{c}_{ij}, \bar{c}_{il}] - E[\bar{c}_{ij}]E[\bar{c}_{il}], \quad j \neq l, \quad l = 1, \dots, n$$

and

$$E[\bar{c}_{ij}, \bar{c}_{il}] = \sum_{s_i=1}^{S_i} p_{is_i} c_{ijs_i} c_{ils_i}.$$

In (7), since

$$\sum_{j=1}^{n} \alpha_{ij} x_j - h_i^1 + h_i^0 > 0.$$

and $\boldsymbol{x}^T V_i \boldsymbol{x} \geq 0$, an optimal solution of the following problem is equivalent to that of (8).

$$\left. \begin{array}{l} \min \ z_i(\boldsymbol{x}) \triangleq \sqrt{V[\Pi_{\bar{Y}_i}(\tilde{G}_i)]}, \ i = 1, \dots, k \\ \text{s. t. } A\boldsymbol{x} \leq \boldsymbol{b}, \ \boldsymbol{x} \geq \boldsymbol{0} \end{array} \right\} \tag{8}$$

In the next section, we consider a method for solving problem (8).

3 Interactive Decision Making Using the Variance Minimization Model Based on a Possibility Measure

Since problem (8) has several objective functions, there does not generally exist the solution optimizing all functions. Therefore, in this section, we discuss the interactive decision making based on the reference point method [13] to obtain a Pareto optimal solution.

For each of the multiple conflicting objective functions, assume that the decision maker can specify the so-called reference point $\bar{\boldsymbol{\pi}} = (\bar{\pi}_1, \dots, \bar{\pi}_k)$ which reflects in some sense the desired values of the objective functions of the decision maker. Also assume that the decision maker can change the reference point interactively due to learning or improved understanding during the solution process. When the decision maker specifies the reference point $\bar{\boldsymbol{\pi}} = (\bar{\pi}_1, \dots, \bar{\pi}_k)$, the corresponding Pareto optimal solution, which is, in the minimax sense, nearest to the reference point or better than that if the reference point is attainable, is obtained by solving the following minimax problem:

$$\left.\begin{array}{l} \min \max\limits_{1 \le i \le k} \{z_i(\boldsymbol{x}) - \bar{\pi}_i\} \\ \text{s. t. } A\boldsymbol{x} \le \boldsymbol{b},\ \boldsymbol{x} \ge \boldsymbol{0} \end{array}\right\} \tag{9}$$

For simplicity, we define N_i and Q_i as

$$z_i(\boldsymbol{x}) - \bar{\pi}_i = \frac{\sqrt{\boldsymbol{x}^t V_i \boldsymbol{x}} - \bar{\pi}_i \left(\sum\limits_{j=1}^{n} \alpha_{ij} x_j - h_i^1 + h_i^0 \right)}{\sum\limits_{j=1}^{n} \alpha_{ij} x_j - h_i^1 + h_i^0} \triangleq \frac{N_i(\boldsymbol{x})}{Q_i(\boldsymbol{x})}$$

Then, in minimax problem (9), the numerators of objective functions are all convex functions and the denominators are all affine functions. Hence, it follows that all objective functions in (8) are quasi-convex functions. Accordingly, the problem is solved by the method of Borde et al.[14], which is an extended version of Dinkelbach-type algorithm.

From the above discussion, an algorithm for obtaining a satisficing solution of a decision maker through interaction is described as follows:

[An interactive satisficing method for fuzzy random multiobjective linear programming problems]

Step 1: Set the initial reference values $\bar{\pi}_i,\ i = 1, \dots, k$ to 0s.
Step 2: Set $\lambda \leftarrow 0$ and find a feasible solution. Let the solution be $\boldsymbol{x}^\lambda$.
Step 3: Calculate q^λ defined by

$$q^\lambda = \max_{1 \le i \le k} \left\{ \frac{N_i(\boldsymbol{x}^\lambda)}{Q_i(\boldsymbol{x}^\lambda)} \right\}$$

and solve the following problem:

$$\begin{array}{l} \min Z \\ \text{s.t. } \dfrac{1}{Q_i(\boldsymbol{x}^\lambda)} \{Q_i(\boldsymbol{x}) - q^\lambda N_i(\boldsymbol{x})\} \le Z,\ i = 1, \dots, k, \\ \boldsymbol{x} \in X. \end{array} \tag{10}$$

Let an optimal solution of (10) be $\boldsymbol{x}^c$. Go to Step 4.
Step 4: If $Z = 0$, then go to Step 5. Otherwise, set $\boldsymbol{x}^\lambda \leftarrow \boldsymbol{x}^c$, $\lambda \leftarrow \lambda + 1$ and return to Step 3.
Step 5: If the decision maker is satisfied with the current solution $\boldsymbol{x}^c$, then terminate. Otherwise, update $\bar{\boldsymbol{\pi}},\ i = 1, \dots, k$ and return to Step 2.

It should be noted that an optimal solution of (9) is at least a weak Pareto optimal solution of (8).

4 Conclusion

In this paper, we have proposed the model to minimize the variances of of degrees of possibility for a multiobjective linear programming problem including fuzzy random variable coefficients. After transforming the formulated problem into the deterministic equivalent multiobjective quasi-convex programming problem, we proposed an interactive satisficing method and considered a solution procedure based on an extended version of Dinkelbach-type algorithm. Although we dealt with only a degree of possibility in this paper, we can also consider the model to minimize the variances of degrees of necessity in similar manner. In future, we will try to consider another model or apply the proposed model to the combinatorial optimization.

References

1. Vajda, S. (1972) Probabilistic Programming, Academic Press.
2. Sakawa, M. (1993) Fuzzy Sets and Interactive Multiobjective Optimization, Plenum, New York.
3. Inuiguchi M., Ramik, J. (2000) Possibilistic linear programming: a brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem, Fuzzy Sets and Systems **111**, 3–28.
4. Kwakernaak, H. (1978) Fuzzy random variable-1: definitions and theorems, Information Sciences **15**, 1–29.
5. Puri, M.L., Ralescu, D.A. (1986) Fuzzy random variables, Journal of Mathematical Analysis and Applications **114**, 409–422.
6. Wang, G.-Y., Zhong, Q. (1993) Linear programming with fuzzy random variable coefficients, Fuzzy Sets and Systems **57**, 295–311.
7. Luhandjula, M.K., Gupta, M.M. (1996) On fuzzy stochastic optimization, Fuzzy Sets and Systems **81**, 47–55.
8. Katagiri, H., Ishii, H. (2000) Chance constrained bottleneck spanning tree problem with fuzzy random edge costs, Journal of the Operations Research Society of Japan **43**, 128–137.
9. Katagiri, H., Sakawa, M., Ishii, H. (2001) Multiobjective fuzzy random linear programming using E-model and possibility measure, Proceedings of Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver, Canada, July 25-28, 2295–2300.
10. Zadeh, L.A., Probability measure of fuzzy events, Journal of Mathematical Analysis and Applications **23**, 421–427.
11. Charnes, A., Cooper, W.W. (1959) Chance constrained programming, Management Science **6**, 73–79.
12. Kaufman, A., Gupta, M.M. (1985) Introduction to Fuzzy Arithmetic: Theory and Applications, Van Nostrand Reinhold Company.
13. Wierzbicki, A.P. (1980) The use of reference objectives in multiobjective optimization, in: G. Frande and T. Gal (eds.), Multiple Criteria Decision Making: Theory and Application, Springer-Verlag.
14. Borde, J., Crouzeix, J.P. (1987) Convergence of a Dinkelbach-type algorithm in generalized fractional programming, Zeitschrift fur Operations Research **31**, 31–54.

On Saddle Points of Multiobjective Functions

Kenji Kimura[1], El Mostafa Kalmoun[2], and Tamaki Tanaka[1]

[1] Graduate School of Science and Technology, Niigata University, Niigata 950-2181, Japan
E-mail: tamaki@m.sc.niigata-u.ac.jp

[2] University of Erlangen, Institute of Computer Science X, D-91058 Erlangen, Germany

Abstract. This paper is concerned with existence theorems for cone-saddle points of vector-valued functions in finite dimensional Euclidean spaces. By means of vector variational-like inequalities, we first characterize a vector saddle point problem and obtain the existence result under some conditions on the subdifferentiable of the vector-valued function. The continuity and convexity assumptions on the objective function are then relaxed.

1 Introduction

Studies on vector-valued minimax theorems or vector saddle point problems have been extended widely; see [9] and references cited therein. Existence results for cone saddle points are based on some fixed point theorems or scalar minimax theorems; see [8]. In 2000, this kind of problems was solved by a different approach in [4], where the connection to a vector variational inequality problem in a finite dimensional vector space was established. In [5], we have extended it into a generalized version under the setting of a normed space, and then presented another generalization to vector problems involving the concept of cone invexity in the general setting of a normed space in The 5th International Conference on Optimization: Techniques and Applications. In this paper, under the restriction to only finite dimensional Euclidean spaces we obtain a new existence result which extends our recent results by relaxing the continuity and convexity assumptions.

2 Preliminary and terminology

Let X and Y be nonempty subsets of finite dimensional spaces R^n and R^m, respectively. Given a vector-valued function $L : X \times Y \to R^p$, the Vector Saddle-Point Problem, (P) for short, is to find $x_0 \in X$ and $y_0 \in Y$ such that

$$L(x_0, y_0) - L(x, y_0) \notin \operatorname{int} R^p_+, \ \forall x \in X, \tag{1a}$$

$$L(x_0, y) - L(x_0, y_0) \notin \operatorname{int} R^p_+, \ \forall y \in Y. \tag{1b}$$

A solution (x_0, y_0) of (P) is called a weak R^p_+-saddle point of function L.

Suppose that η is a vector-valued function from $X \times X$ to R^n.

Definition 1. For any given R^p_+-convex function $f : X \to R^p$, the vector subdifferential of f at $a \in X$ with respect to η is a set of linear operators from R^n to R^p denoted by

$$\partial f(a) := \{ A \in \mathcal{L}(R^n, R^p) \mid f(x) - f(a) - \langle A, \eta(x,a) \rangle \in R^p_+ \ \forall x \in X \}. \quad (2)$$

If $\partial f(a)$ is nonempty for every $a \in X$ then f is subdifferentiable in X.

For each $x_0 \in X$, we define a maximal solution set of $L(x_0, Y)$ with respect to $\operatorname{int} R^p_+$ as follows:

$$T(x_0) := \{ y \in Y \mid L(x_0, v) - L(x_0, y) \notin \operatorname{int} R^p_+, \ \forall v \in Y \}. \quad (3)$$

For each $x_0 \in X$, it follows from the nonemptyness and compactness of $\cup_{v \in Y} L(x_0, v)$ and the correctness of R^p_+ that $T(x_0)$ is nonempty; see Theorem 2.6 in [5]. Letting $f(\cdot) = L(\cdot, y_0)$ for a fixed y_0 in (2), we consider the vector subdifferential $\partial L(\cdot, y_0)$ and then the following Vector Variational-like Inequality Problem, (Q) for short, is to find $x_0 \in X$ and $y_0 \in T(x_0)$ such that

$$\langle A, \eta(x, x_0) \rangle \notin -\operatorname{int} R^p_+, \ \forall x \in X \text{ for some } A \in \partial L(x_0, y_0). \quad (4)$$

Definition 2. A multifunction $F : X \to 2^{R^p}$ is called upper-semicontinuous, u.s.c. for short, if for every $x \in X$ and $U_{F(x)} \subset R^p$, a neighborhood of $F(x)$, there exists $V_x \subset X$, a neighborhood of x such that $F(y) \subset U_{F(x)}$ for all $y \in V_x$.

Definition 3. Let $\mathcal{X}$ and $\mathcal{Y}$ be two metric spaces. A set-valued map $F : \mathcal{X} \to 2^{\mathcal{Y}}$ is said to be closed at x_0 if for any sequences $\{x_n\}$ with $x_n \to x_0$ and $\{y_n\}$ with $y_n \in F(x_n)$, $y_n \to y_0$ for some $y_0 \in Y$ implies that $y_0 \in F(x_0)$.

Theorem 1. (See Theorem 2.3 in [5]) *For each $x_0 \in X$, $T(x_0)$ is closed.*

3 Existense results of cone saddle points

Theorem 2. *Suppose that X is convex and L is vector subdifferentiable with respect to η in the first argument. Then the solution set of* (Q) *is included in that of* (P).

Theorem 3. (**Fan-KKM Theorem**, see Lemma 1 in [2]) *Let Y be a subset of the topological vector space X. For each $x \in Y$, let a closed set $F(x)$ in X be given such that $F(x)$ is compact for at least one $x \in Y$. If the convex hull of every finite subset $\{x_1, \ldots, x_n\}$ of Y is contained in the corresponding union $\bigcup_{i=1}^n F(x_i)$, then $\bigcap_{x \in Y} F(x) \neq \phi$.*

Based on Theorem 3, we obtain our existence results. The mapping $F : Y \to 2^Y$ is called the KKM-map if $\operatorname{conv}\{x_1, \ldots, x_n\} \subset \bigcup_{i=1}^n F(x_i)$ for every finite subset $\{x_1, \ldots, x_n\}$ of Y, where $\operatorname{conv} D$ denotes the convex hull of the set D.

Theorem 4. *Let X and Y be nonempty closed convex subset and nonempty compact in R^n and R^m, respectively. Assume that the vector-valued function L is subdifferentiable with respect to η in the first argument, where $\eta : X \times X \to R^n$ satisfies the following three conditions: for all $x \in X$,*

(i) $\eta(\cdot, x)$ *is affine,*
(ii) $\eta(x, \cdot)$ *is continuous, and*
(iii) $\eta(x, x) = 0$.

Assume that $\partial L(x, y)$ is u.s.c. in both x and y. If there exist a nonempty compact subset B of R^p and $x_0 \in (B \cap X)$ such that for any $x \in (X \setminus B)$, $y \in T(x)$, $A \in \partial L(x, y)$

$$\langle A, \eta(x_0, x)\rangle \in -\text{int}\, R^p_+,$$

then problem (P) *has at least one solution.*

Proof. Define a multifunction $F : X \to 2^X$ by

$$F(u) := \{\, x \in X \mid \langle A, \eta(u, x)\rangle \notin -\text{int}\, R^p_+, \\ \text{for some } y \in T(x) \text{ and } A \in \partial L(x, y)\,\}, \quad u \in X.$$

In order to prove the theorem, it is sufficient to show by Theorem 2 that problem (Q) has at least one solution pair $(x_0, y_0) \in X \times T(x_0)$. So we should show, by Fan-KKM Theorem, the following three points:

(a) F is a KKM-map;
(b) $F(x)$ is closed for each $x \in X$; and
(c) there exists $\hat{x} \in X$ such that $F(\hat{x})$ is compact.

First, we prove condition (1). Suppose to the contrary that there exist $x_1, x_2, \ldots, x_m$ and $\alpha_1, \alpha_2, \ldots, \alpha_m$ such that

$$\hat{x} := \sum_{i=1}^{m} \alpha_i x_i \notin \bigcup_{i=1}^{m} F(x_i), \quad \sum_{i=1}^{m} \alpha_i = 1.$$

Then, $\hat{x} \notin F(x_i)$ for all $i = 1, \ldots, m$, and hence for any $y \in T(\hat{x})$, $A \in \partial L(\hat{x}, y)$,

$$\langle A, \eta(x_i, \hat{x})\rangle \in -\text{int}\, R^p_+,$$

for all $i = 1, \ldots, m$. Since $\text{int}\, R^p_+$ is convex, we have

$$\sum_{i=1}^{m} \alpha_i \langle A, \eta(x_i, \hat{x})\rangle \in -\text{int} R^p_+.$$

Since A is a linear operater and η is an affine operater, we have

$$\left\langle A, \eta\left(\sum_{i=1}^{m} \alpha_i x_i, \sum_{i=1}^{m} \alpha_i \hat{x}\right)\right\rangle \in -\text{int}\, R^p_+.$$

Therefore

$$\langle A, \eta(\hat{x}, \hat{x})\rangle = 0 \in -\operatorname{int} R^p_+,$$

which is inconsistent. Thus, we deduce that

$$\operatorname{conv}\{x_1, x_2, \ldots, x_m\} \subset \bigcup_{i=1}^{m} F(x_i).$$

Next, we show that the condition (b) holds. For each $u \in X$, let $\{x_n\} \subset F(u)$ such that $x_n \to x \in X$. Since $x_n \in F(u)$ for all n, there exist $y_n \in T(x_n)$ and $A_n \in \partial L(x_n, y_n)$ such that

$$\langle A_n, \eta(u, x_n)\rangle \in W,$$

where $W := R^p \backslash (-\operatorname{int} R^p_+)$. As $\{y_n\} \subset Y$, without loss of generality, we can assume that there exists $y \in Y$ such that $y_n \to y$. Now T is closed, by the reason of Theorem 1, so $y \in T(x)$. Because of the closedness of W, the upper semicontinuity of ∂L and $\langle A_n, \eta(u, x_n)\rangle \in W$ for all n, there exists $A \in \partial L(x, y)$ such that

$$\langle A, \eta(u, x)\rangle \in W.$$

Hence $x \in F(u)$. As a result the condition (b) holds.

Finally in order to prove the condition (c). Since $F(\bar{x})$ is closed and B is compact, it is sufficient to show that $F(\bar{x}) \subset B$. Suppose to the contrary that there exists $\hat{x} \in F(\bar{x})$ such that $\hat{x} \notin B$. Since $\hat{x} \in F(\bar{x})$, there exist $\hat{y} \in T(\hat{x})$ and $\hat{A} \in \partial L(\hat{x}, \hat{y})$ such that

$$\langle \hat{A}, \eta(\bar{x}, \hat{x})\rangle \notin -\operatorname{int} R^p_+. \tag{5}$$

Since $\hat{x} \notin B$, by the hypothesis, for any $y \in T(\hat{x})$ and $A \in \partial L(\hat{x}, y)$,

$$\langle A, \eta(\bar{x}, \hat{x})\rangle \in -\operatorname{int} R^p_+,$$

which contradicts condition (1). Hence $F(\bar{x}) \subset B$. Since B is compact and $F(\bar{x})$ is also closed, $F(\bar{x})$ is compact, that is the condition (c) holds. Consequently by Fan-KKM Theorem, it follows that $\bigcap_{x \in X} F(x) \neq \phi$. Thus, there exists $x_0 \in X$ and $y_0 \in T(y_0)$ such that

$$\langle A, \eta(x, x_0)\rangle \notin -\operatorname{int} R^p_+,$$

for all $x \in X$. As a result there exists at least one solution of (P).

Definition 4. Suppose we are given vector-valued functions f and h, which consist of p real-valued functions $f_1, \ldots, f_p$ and $h_1, \ldots, h_p$ on $X \times Y$, respectively. h is said to be a vector convex envelope of f if h_i is the convex envelope of f_i for every $i \in \{1, \ldots, p\}$.

Assuption A. For $f : X \to R^p$ and its vector convex envelope h, the following condition holds:

$$\{ x \in X \mid h(x) - h(y) \notin \mathrm{int} R^p_+ \ \forall y \in X \} \\ \subset \{ x \in X \mid f(x) - f(y) \notin \mathrm{int} R^p_+ \ \forall y \in X \}.$$

Corollary 1. *Let X and Y be nonempty closed convex and nonempty compact in R^n and R^m, respectively. Suppose that a vector-valued function $H : X \times Y \to R^p$ is the convex envelope of $L : X \times Y \to R^p$ in the first argument and that H satisfies the conditions on L in Theorem 4. If $h(x) := H(x, y)$ and $f(x) := L(x, y)$ satisfy Assumption A for each $y \in Y$ and L is continuous with respect to the second argument, then problem* (P) *has at least one solution.*

Proof. In the definition of F in the proof of Theorem 4, we replace $\partial L(x, y)$ by $\partial H(x, y)$. Then we see there exist $x_0 \in X$, $y_0 \in T(x_0)$ and $A \in \partial H(x_0, y_0)$ such that

$$\langle A, \eta(x, x_0) \rangle \notin -\mathrm{int}\, R^p_+ \quad \forall x \in X.$$

Assumption A leads to

$$L(x_0, y_0) - L(x, y_0) \notin \mathrm{int}\, R^p_+, \quad \forall x \in X$$

and $y_0 \in T(x_0)$ leads to

$$L(x_0, y) - L(x_0, y_0) \notin \mathrm{int}\, R^p_+, \quad \forall y \in Y.$$

Which means there exists at least one solution of (P).

References

1. G. Y. Chen. (1992) Existence of Solutions for a Vector Variational Inequality: An Extension of the Hartmann-Stampacchia Theorem, Journal of Optimization Theory and Applications, Vol.74, pp.445–456.
2. K. Fan. (1961) A Generalization of Tychonoff's Fixed Point Theorem, Mathematische Annalen, Vol.142, pp.305–310.
3. E. M. Kalmoun, H. Riahi, and T. Tanaka. (2001) On Vector Equilibrium Problems: Remarks on a General Existence Theorem and Applications, Nihonkai Mathematical Journal, Vol.12, No.2, pp.149–164.
4. K. R. Kazmi and S. Khan. (2000) Existence of Solutions for a Vector Saddle Point Problem, Bulletin of the Australian Mathmatical Society, Vol.61, pp.201–206.
5. K. Kimura and T. Tanaka. (2001) Existence Theorems of Saddle Points for Vector Valued Functions, to appear in the Proceedings of the Second International Conference on Nonlinear Analysis and Convex Analysis, Yokohama publishers, Tokyo.
6. D. T. Luc. (1989) An Existence Theorem in Vector Optimization, Mathematics of Operations Research, Vol.14, pp.693–699.

7. T. Tanaka, (1994) Generalized quasiconvexities, cone saddle points, and minimax theorem for vector-valued functions, Journal of Optimization Theory and Applications, Vol.81, No.2, pp.355–377.
8. T. Tanaka. (1997) Generalized Semicontinuity and Existence Theorems for Cone Saddle Points, Applied Mathematics and Optimization, Vol.36, pp.313–322.
9. T. Tanaka. (2000) Vector-Valued Minimax Theorems in Multicriteria Games, pp.75–99 (Chapter 5) in "New Frontiers of Decision Making for the Information Technology Era," edited by Yong Shi and Milan Zeleny, World Scientific, Singapore.

An Application of Fuzzy Multiobjective Programming to Fuzzy AHP

Hiroaki Kuwano

Kanazawa Gakuin University, Kanazawa Ishikawa 920-1392, Japan

Abstract. In the real world, we encounter many subjective evaluation problems. For some class of them, we can give the numerical evaluation by means of Analytic Hierarchy Process (AHP) approaches. Moreover, we can even identify the λ-fuzzy measure as the subjective evaluation by using AHP result. Although the original problem includes subjective estimates, this identification process has no vagueness. So, in this article, we discuss the identification process with vagueness. That is, fuzzy multiobjective programming techniques are applied to an identification problem for the parameter of λ-fuzzy measure.

1 Introduction

On a decision making problem, we select the most appropriate alternative among all admissible ones based on an evaluation criterion. In the case that the criterion and alternatives are presented by mathematical models, it is effective for us to approach them using mathematical programming techniques. However, it is not rare that there are plural evaluation criteria on a problem in the real world. In addition, it is even improper to express the criteria by mathematical expressions, that is, subjective evaluation criterion. For example, design and sense. For some class of subjective evaluation problems, we can give the numerical evaluation by means of Analytic Hierarchy Process(AHP) developed by T.L. Saaty(e.g. [4]). By applying AHP, we are able to evaluate the alternatives based on the decision maker's intuitive and/or experiential judgment, and derive a numerical evaluation values for each alternative.

Fuzzy measure and fuzzy integrals(e.g. [6]) are is also known as models of the subjective evaluation. λ-fuzzy measures [5] are especially useful for evaluating alternatives with a sort of mutual relations. However, it is difficult to decide a parameter of the λ-fuzzy measure. For the sake of overcoming such difficulty, we are able to utilize AHP techniques. In other words, we can decide the parameter of the λ-fuzzy measure as the subjective evaluation by using AHP techniques. Then, the λ-fuzzy measure is identified, and assigns a numerical evaluation value for any subset of the set containing all alternatives.

In this article, we discuss an extension of the above method for a subjective evaluation on the power set of all alternatives. Namely, we introduce a subjective evaluation function (called λ-fuzzy measure type evaluation function) with a fuzzy parameter, and formulate the identification problem as a fuzzy multiobjective programming problem by using the results of AHP.

Therefore, the evaluation values are given as fuzzy numbers. Then, decision maker obtains his own subjective evaluation for each subset of the set of all alternatives, i.e. his own subconscious evaluation for mutual relations among all alternatives.

2 Preliminaries

In this section, we recall basic definitions and properties concerned with λ-fuzzy measures, fuzzy sets and the index of ranking fuzzy numbers based on Possibility Theory. Throughout this article, $N = \{1, 2, \ldots, n\}$ denotes a set of all alternatives, and 2^N is the power set of N.

2.1 λ-fuzzy measures

For a fixed parameter $-1 < \lambda < \infty$, a function g_λ defined on 2^N to the unit interval $[0,1]$ is called a λ-fuzzy measure (Sugeno measure) , if and only if it satisfies $g_\lambda(N) = 1$, $g_\lambda(\emptyset) = 0$ and

$$g_\lambda\left(\bigcup_{j=1}^{n} E_j\right) = \begin{cases} \dfrac{1}{\lambda}\left\{\prod_{j=1}^{n}(1+\lambda\cdot g_\lambda(E_j)) - 1\right\}, & \text{if } \lambda \neq 0, \\ \sum_{j=1}^{n} g_\lambda(E_j), & \text{if } \lambda = 0. \end{cases}$$

where $\{E_1, \cdots, E_n\}$ is any family of disjoint subsets of N. As well-known, if $E, F \in 2^N$ are disjoint sets then $g_\lambda(E \cup F) = g_\lambda(E) + g_\lambda(F) + \lambda \cdot g_\lambda(E) \cdot g_\lambda(F)$. Concerning λ-fuzzy measures you can find further details in [5,6], for example.

2.2 Fuzzy sets

$\widetilde{a}$ denotes a fuzzy set on an m-dimensional Euclidean space $\mathbb{R}^m$ with its membership function $\mu_{\widetilde{a}} : \mathbb{R}^m \to [0,1]$. (For details, refer to [3,7]) The α-level set, $0 \leq \alpha \leq 1$, of $\widetilde{a}$ is defined as $[\widetilde{a}]^\alpha = \{x \in \mathbb{R} \mid \mu_{\widetilde{a}}(x) \geq \alpha\}$ for $\alpha \in (0,1]$, and $[\widetilde{a}]^0 = \mathrm{cl}\,(\cup_{0<\alpha\leq 1} [\widetilde{a}]^\alpha)$, where "cl" denotes the closure of the set. A fuzzy set $\widetilde{a}$ on $\mathbb{R}$ is a symmetric triangular fuzzy number, if and only if the membership function is defined by $\mu_{\widetilde{a}}(x) = \max\{|x-a|/\sigma_a, 0\}$, where a is the center, $\sigma_a > 0$ is the spread and we write $\widetilde{a} = (a, \sigma_a)_T$.

According to Dubois and Prade's results [1], for $\widetilde{a} = (a, \sigma_a)_T$, $\widetilde{b} = (b, \sigma_b)_T$ and a scalar $\mu \in \mathbb{R}$, the following are valid:

$$\nu\widetilde{a} = (\nu a, \nu\sigma_a)_T\,, \tag{1a}$$

$$\widetilde{a} \pm \widetilde{b} = (a \pm b, \sigma_a + \sigma_b)_T\,, \tag{1b}$$

$$\widetilde{a} \cdot \widetilde{b} \simeq (ab, |a|\sigma_b + |b|\sigma_a)_T\,, \tag{1c}$$

$$1/\widetilde{a} \simeq \left(1/a, \sigma_a/a^2\right)_T, \quad (a \neq 0). \tag{1d}$$

where "$\simeq$" denotes an approximate equation.

2.3 Ranking fuzzy numbers

Possibility measure([8]) is defined by $\Pi_{\widetilde{b}}(\widetilde{c}) = \sup_x \min\left\{\mu_{\widetilde{c}}(x), \mu_{\widetilde{b}}(x)\right\}$, where $\widetilde{b}$ is a fuzzy number and $\widetilde{c}$ is an arbitrary fuzzy set on $\mathbb{R}$. Applying this concept and necessity measure, Dobuis and Prade proposed four indices of ranking fuzzy numbers based on Possibility Theory in [2]. The following is one of them:

$$\mathrm{Pos}\left(\widetilde{a} \leq \widetilde{b}\right) = \Pi_{\widetilde{b}}([\widetilde{a}, \infty)) = \sup_x \min\left\{\mu_{[\widetilde{a},\infty)}(x), \mu_{\widetilde{b}}(x)\right\}, \tag{2}$$

where a fuzzy interval $[\widetilde{a}, \infty)$ is characterized by the membership function $\mu_{[\widetilde{a},\infty)}(y) = \sup_{x:x\leq y} \mu_{\widetilde{a}}(x)$. They explained that $\mathrm{Pos}\left(\widetilde{a} \leq \widetilde{b}\right)$ yields the grade of possibility of "$\widetilde{a} \leq \widetilde{b}$". Besides, they defined the grade of possibility of "$\widetilde{a} = \widetilde{b}$";

$$\mathrm{Pos}\left(\widetilde{a} = \widetilde{b}\right) = \min\left\{\mathrm{Pos}\left(\widetilde{a} \leq \widetilde{b}\right), \; \mathrm{Pos}\left(\widetilde{a} \geq \widetilde{b}\right)\right\}. \tag{3}$$

By the above Equations (2) and (3), the next assertion is valid.

Proposition 1. *Let $\widetilde{a}$ be a triangular fuzzy number, and let b be a real number. Then, the condition* $\mathrm{Pos}(\widetilde{a} = b) \geq \alpha$, $(0 < \alpha \leq 1)$ *is equivalent to* $b \in [\,\widetilde{a}\,]^{\alpha}$.

3 Subjective evaluation

First of all, we should make sure of our purpose. It is to propose a subjective evaluation method based on AHP and λ-fuzzy measure. The evaluated objects are subsets of the set of all alternatives, assuming that there are some mutual relations among them and the decision maker perceives the interaction but his perception is not clear.

Our proposing method is constructed from two steps. In Step 1, the decision maker applies AHP techniques to his problem. In Step 2, he identifies his own subjective evaluation function.

Step 1.

The decision maker evaluates $S \in 2^N$ by AHP in advance, i.e. calculates the importance for every nonempty subset S. In the rest of this article, the pairwise comparison matrix $A = (a_{ST}) \in \mathbb{R}^{(2^n-1)\times(2^n-1)}$ and the vector

$\boldsymbol{w} = (w_S) \in \mathbb{R}^{(2^n-1)}$ denote

$$A = \begin{pmatrix} a_{\{1\},\{1\}} & a_{\{1\},\{2\}} & \cdots & a_{\{1\},\{1,2\}} & \cdots & a_{\{1\},N} \\ a_{\{2\},\{1\}} & a_{\{2\},\{2\}} & \cdots & a_{\{2\},\{1,2\}} & \cdots & a_{\{2\},N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ a_{\{1,2\},\{1\}} & a_{\{1,2\},\{2\}} & \cdots & a_{\{1,2\},\{1,2\}} & \cdots & a_{\{1,2\},N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ a_{N,\{1\}} & a_{N,\{2\}} & \cdots & a_{N,\{1,2\}} & \cdots & a_{N,N} \end{pmatrix},$$

$$\boldsymbol{w} = \left(w_{\{1\}}\ w_{\{2\}}\ \dots\ w_{\{1,2\}}\ \dots\ w_N\right),$$

respectively.

For our purpose, we should modify λ-fuzzy measure as follows.

Definition 1. Let $0 \le \alpha \le 1$ be a fixed grade, And let g be a fuzzy set-valued map on 2^N, i.e. $g : 2^N \to \mathcal{F}(\mathbb{R})$, where $\mathcal{F}(\mathbb{R})$ denotes the set of all fuzzy set on $\mathbb{R}$. Then, g is a λ-fuzzy measure type evaluation function with a fuzzy parameter $\widetilde{\lambda} > 0$, if g satisfies the following conditions:

1. $g(\emptyset) = 0$,
2. $g(\{i\}) = w_{\{i\}}$ for all $i \in N$,
3. $g(S) = \frac{1}{\widetilde{\lambda}} \left\{\prod_{i \in S} \left(1 + \widetilde{\lambda} \cdot g(\{i\})\right) - 1\right\}$ for all $S \in \mathcal{S}$,
4. $\mathrm{Pos}(g(S) = w_S) \ge \alpha$ for all $S \in \mathcal{S}$,

where $\mathcal{S}$ denotes $2^N \setminus \{\emptyset, \{1\}, \{2\}, \dots, \{n\}\}$.

We write the above g by $g_{\widetilde{\lambda},\alpha}$ or simply $g_{\widetilde{\lambda}}$ in order to exhibit parameters.

Remark 1. Properly speaking, $g(S)$ should be calculated by Zadeh's extension principle for each $S \in \mathcal{S}$ in Definition 1. However, this way gives rise to computational trouble. So, we may calculate them by composition of four operations (1a), (1b), (1c) and (1d).

Step 2.

Based on results of Step 1, the decision maker identifies his λ-fuzzy measure type evaluation function $g_{\widetilde{\lambda}}$. Then, if we restrict the parameter $\widetilde{\lambda}$ with symmetric triangular fuzzy numbers, images of the evaluation function are also approximate symmetric triangular fuzzy numbers by above mentioned Dubois and Prade's results. So, we assume that he selects an appropriate fuzzy parameter $\widetilde{\lambda} = (\lambda, \sigma)_T > 0$. Therefore, he has the following problem:

$$\begin{array}{ll} \text{Find} & \widetilde{\lambda} = (\lambda, \sigma)_T \\ \text{such that} & g_{\widetilde{\lambda}}(\emptyset) = 0, \\ & g_{\widetilde{\lambda}}(\{i\}) = w_{\{i\}},\ \forall i \in N, \\ & g_{\widetilde{\lambda}}(S) = \frac{1}{\widetilde{\lambda}} \left\{\prod_{i \in S} \left(1 + \widetilde{\lambda} \cdot g_{\widetilde{\lambda}}(\{i\})\right) - 1\right\},\ \forall S \in \mathcal{S}, \\ & \mathrm{Pos}\left(g_{\widetilde{\lambda}}(S) = w_S\right) \ge \alpha,\ \forall S \in \mathcal{S}, \\ & \widetilde{\lambda} = (\lambda, \sigma)_T > 0. \end{array}$$

The above problem is able to formulated as a substitutive fuzzy multiobjective programming problem

$$\begin{aligned} &\text{maximize}_{\lambda,\sigma}\ \mathrm{Pos}\big(w_S = g_{\tilde{\lambda}}(S)\big)\,, \quad S \in \mathcal{S}, \\ &\text{maximize}_{\lambda,\sigma}\ \frac{\lambda}{\sigma}\,, \\ &\text{subject to} \quad \tilde{\lambda} = (\lambda, \sigma)_T > 0. \end{aligned} \tag{P$_1$}$$

Problem P_1 is equivalent to the following problem by Proposition 1.

$$\begin{aligned} &\text{maximize}\ \alpha, \\ &\text{maximize}\ \frac{\lambda}{\sigma}\,, \\ &\text{subject to}\ w_S \in \left[\, g_{\tilde{\lambda}}(S)\, \right]^{\alpha}, \quad S \in 2^N, \\ &\qquad\qquad \tilde{\lambda} = (\lambda, \sigma)_T > 0. \end{aligned} \tag{P$_2$}$$

Obviously, the above fuzzy multiobjective problem P_1 (or P_2) has some Pareto optimal solutions, however they may be not always reasonable.

Avoiding irrationality, we assume that the decision maker gives an aspiration level $0 < \alpha \leq 1$. Then, his problem is the following.

$$\begin{aligned} &\text{maximize}\ \frac{\lambda}{\sigma}\,, \\ &\text{subject to}\ w_S \in \left[\, g_{\tilde{\lambda}}(S)\, \right]^{\alpha}, \quad S \in 2^N, \\ &\qquad\qquad \lambda > \sigma > 0. \end{aligned} \tag{P$_3$}$$

If there exists an optimal solution (λ^*, σ^*) for P_3, then the triplet $(\lambda^*, \sigma^*, \alpha)$ is a Pareto optimal solution for P_2. Then, by setting $\tilde{\lambda}(\alpha) = (\lambda^*, \sigma^*)_T$, we obtain that the decision maker's λ-fuzzy measure type evaluation function $g_{\tilde{\lambda}(\alpha)}$.

4 A numerical Example

To illustrate our subjective evaluation method, we consider the set of all alternatives $N = \{1, 2, 3\}$. Then $\mathcal{S} = \{\{1,2\}, \{2,3\}, \{1,3\}, N\} \subset 2^N$. Suppose that the decision maker decides his pairwise comparison matrix

$$A = \begin{pmatrix} 1 & 1/7 & 1/7 & 1/9 & 1/9 & 1/7 & 1/9 \\ 7 & 1 & 1 & 1/3 & 1/7 & 1 & 1/5 \\ 7 & 1 & 1 & 1/3 & 1/3 & 1/5 & 1/5 \\ 9 & 3 & 3 & 1 & 1/5 & 1 & 1/5 \\ 9 & 7 & 3 & 5 & 1 & 5 & 1/7 \\ 7 & 1 & 5 & 1 & 1/5 & 1 & 1/5 \\ 9 & 5 & 6 & 5 & 7 & 5 & 1 \end{pmatrix}.$$

with two levels hierarchical structure constructed from a criterion and alternatives. Then, we get the following normalized importance vector:

$$\boldsymbol{w} = \big(0.035\ 0.124\ 0.117\ 0.211\ 0.513\ 0.201\ 1.000\big)\,.$$

Next we calculate $g_{\tilde{\lambda}}(S)$ for each $S \in \mathcal{S}$.

$$g_{\tilde{\lambda}}(\{1,2\}) = \left(0.159 + 0.004\lambda, \frac{\lambda}{\sigma}(0.319 + 0.013\lambda)\right)_T,$$
$$g_{\tilde{\lambda}}(\{2,3\}) = \left(0.241 + 0.015\lambda, \frac{\lambda}{\sigma}(0.482 + 0.044\lambda)\right)_T,$$
$$g_{\tilde{\lambda}}(\{1,3\}) = \left(0.152 + 0.004\lambda, \frac{\lambda}{\sigma}(0.303 + 0.012\lambda)\right)_T,$$
$$g_{\tilde{\lambda}}(N) = \left(0.276 + 0.023\lambda + 0.001\lambda^2, \frac{\lambda}{\sigma}(0.552 + 0.069\lambda + 0.002\lambda^2)\right)_T$$

Now, suppose that the decision maker gives an aspiration level $\alpha = 0.7$. We formulate the identification problem:

$$\begin{aligned}
\max\ & \frac{\lambda}{\sigma}, \\
\text{s.t. } & 0.211 \in \left[\left(0.159 + 0.004\lambda, \tfrac{\lambda}{\sigma}(0.319 + 0.013\lambda)\right)_T\right]^{0.7}, \\
& 0.513 \in \left[\left(0.241 + 0.015\lambda, \tfrac{\lambda}{\sigma}(0.482 + 0.044\lambda)\right)_T\right]^{0.7}, \\
& 0.201 \in \left[\left(0.152 + 0.004\lambda, \tfrac{\lambda}{\sigma}(0.303 + 0.012\lambda)\right)_T\right]^{0.7}, \\
& 1.000 \in \left[\left(0.276 + 0.023\lambda + 0.001\lambda^2, \tfrac{\lambda}{\sigma}(0.552 + 0.069\lambda + 0.002\lambda^2)\right)_T\right]^{0.7}, \\
& \lambda > \sigma > 0.
\end{aligned}$$

This problem has an optimal solution $(\lambda^*, \sigma^*) = (18.493, 3.104)$. Therefore, the decision maker obtains his own subjective evaluation for each subset of the set of all alternatives as following:

$$\begin{aligned}
g_{\tilde{\lambda}(\alpha)}(\emptyset) &= 0, & g_{\tilde{\lambda}(\alpha)}(\{1\}) &= 0.035, \\
g_{\tilde{\lambda}(\alpha)}(\{2\}) &= 0.124, & g_{\tilde{\lambda}(\alpha)}(\{3\}) &= 0.117, \\
g_{\tilde{\lambda}(\alpha)}(\{1,2\}) &= (0.240, 0.094)_T, & g_{\tilde{\lambda}(\alpha)}(\{2,3\}) &= (0.509, 0.216)_T, \\
g_{\tilde{\lambda}(\alpha)}(\{1,3\}) &= (0.227, 0.089)_T, & g_{\tilde{\lambda}(\alpha)}(N) &= (0.873, 0.423)_T.
\end{aligned}$$

5 Conclusions

In this article, we introduced a fuzzy evaluation system on the power set. In the process of identification of the decision maker's evaluation function, a fuzzy multiobjective programming problem is used. By this evaluation function, the decision maker can evaluate the objects subjectively.

Acknowledgements

The author is grateful to the referee for suggesting improvements and their valuable comments.

References

1. Dubois, D., Prade, H. (1978) Operations on Fuzzy Numbers, Int. J. Systems Sci. **9**, 613–626
2. Dubois, D., Prade, H. (1983) Ranking Fuzzy Numbers in the Setting of Possibility Theory, Inform. Sci. **30**, 183–224
3. Lowen, R., (1996) Fuzzy Set Theory, Kluwer Academic Publishers, Netherlands
4. Saaty, T. L. (1980) The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation, McGraw-Hill, New York
5. Sugeno, M. (1974) Theory of Fuzzy Integrals and its Applications. Ph.D. dissertation, Tokyo Institute of Technology.
6. Wang, Z., Klir, G. J. (1992) Fuzzy Measure Theory, Plenum Press, New York
7. Zadeh, L. A. (1965) Fuzzy Sets, Inform. and Control **8**, 338–353
8. — (1978) Fuzzy Sets as a Basis for a Theory of Possibility, Fuzzy Sets and Systems **1**, 3–28

On Affine Vector Variational Inequality

Gue Myung Lee and Kwang Baik Lee

Department of Applied Mathematics, Pukyong National University, Pusan 608-737, Korea

Abstract. The concept of vector variational inequality was introduced by Giannessi ([3]) in 1980. Since then, various vector variational inequalities and their applications to multiobjective (vector) optimization problems have been studied. Very recently, many authors ([1,4,9,10,11]) have investigated the connectedness of solution sets of vector variational inequalities.

In this paper, we study the connectedness of solution sets for affine vector variational inequalities with 2×2 monotone matrices. Moreover, we give examples to clarify our result on the connectedness.

Key words and phrases: affine vector variational inequality, boundedness, connectedness.

1 Introduction and Preliminaries

Let $\Lambda = \{\xi = (\xi_1, \xi_2) \in I\!R^2 \mid \xi_i \geqq 0,\ i = 1, 2,\ \sum_{i=1}^2 \xi_i = 1\}$, and $\overset{o}{\Lambda} = \{\xi = (\xi_1, \xi_2) \in I\!R^2 \mid \xi_i > 0,\ i = 1, 2,\ \sum_{i=1}^2 \xi_i = 1\}$. Let $\triangle = \{x \in I\!R^2 \mid Ax \geqq b\}$, where $A \in I\!R^{m\times 2}$ and $b \in I\!R^m$. Let $\langle \cdot, \cdot \rangle$ denote the inner product on $I\!R^2$.

Assume that $\triangle \neq \emptyset$. Let $M_i \in I\!R^{2\times 2}$ and $q_i \in I\!R^2$, $i = 1, 2$.

Consider the following affine vector variational inequalities:

(VVI) Find $\bar{x} \in \triangle$ such that
$$(\langle M_1\bar{x} + q_1,\ x - \bar{x}\rangle, \langle M_2\bar{x} + q_2,\ x - \bar{x}\rangle) \notin -I\!R^2_+ \setminus \{0\} \quad \forall x \in \triangle,$$

(VVI)^w Find $\bar{x} \in \triangle$ such that
$$(\langle M_1\bar{x} + q_1,\ x - \bar{x}\rangle, \langle M_2\bar{x} + q_2,\ x - \bar{x}\rangle) \notin -\text{int}I\!R^2_+ \quad \forall x \in \triangle,$$

where $I\!R^2_+ = \{x := (x_1, x_2) \in I\!R^2 \mid x_i \geqq 0,\ i = 1, 2\}$ and $\text{int}I\!R^2_+$ is the interior of $I\!R^2_+$, and consider their related scalar variational inequality: let $\xi = (\xi_1, \xi_2) \in \Lambda$.

(VI)_ξ Find $\bar{x} \in \triangle$ such that
$$\left\langle \sum_{i=1}^2 \xi_i M_i \bar{x} + \sum_{i=1}^2 \xi_i q_i,\ x - \bar{x} \right\rangle \geqq 0 \quad \forall x \in \triangle.$$

We denote the solution sets of (VVI), $(\text{VVI})^w$ and $(\text{VI})_\xi$ by sol(VVI), $\text{sol}(\text{VVI})^w$ and $\text{sol}(\text{VI})_\xi$, respectively.

It is clear that $\text{sol(VVI)} \subset \text{sol}(\text{VVI})^w$.

From Theorem 2.1 in (Lee, Kim, Lee and Yen [4]) and Theorem 2.1 in (Lee, Yen [5]), we can obtain the following proposition:

Proposition 1.

$$sol(VVI) = \bigcup_{\xi \in \overset{o}{\Lambda}} sol(VI)_\xi \subset sol(VVI)^w = \bigcup_{\xi \in \Lambda} sol(VI)_\xi.$$

Now we give some well-known results for multifunctions, which will be used for the proof of our main result.

Let X, Y be two topological spaces and $G : X \to 2^Y$ a multifunction.

Definition 1. The space X is said to be connected if there do not exist nonempty open subsets $V_i \subset X$, $i = 1, 2$, such that

$$V_1 \cap V_2 = \emptyset \quad \text{and} \quad V_1 \cup V_2 = X.$$

Definition 2. (i) The multifunction G is said to be closed if its graph, $\{(x, y) \in X \times Y \mid y \in G(x)\}$, is closed in $X \times Y$.
(ii) The multifunction G is said to be upper semicontinuous (shortly u.s.c.) if for every $a \in X$ and every open set $\Omega \subset Y$ satisfying $G(a) \subset \Omega$, there exists a neighborhood U of a such that $G(a') \subset \Omega \quad \forall a' \in U$.

Lemma 1. (Warburton [8], Theorem 3.1) *Assume that X is connected. If for every $x \in X$, the set $G(x)$ is nonempty and connected, and G is upper semicontinuous, then the set $G(X) := \cup_{x \in X} G(x)$ is connected.*

In general, if $\text{sol}(\text{VVI})^w$ is bounded, then sol(VVI) and $\text{sol}(\text{VVI})^w$ are connected. However, the boundedness of sol(VVI) may not imply the boundedness of $\text{sol}(\text{VVI})^w$ (see Example 2.1). So, we can raise one question: when sol(VVI) is bounded, are sol(VVI) and $\text{sol}(\text{VVI})^w$ connected ?

In this paper, we show that we can give a positive answer for the question about affine vector variational inequalities for monotone 2×2 matrices. Furthermore we give examples to clarify our result on the connectedness.

2 Main Result

Now we give our main result:

Theorem 1. *Suppose that M_1, M_2 are monotone on $\triangle$, that is, for each $i = 1, 2$, $\langle x - y,\ M_i(x - y)\rangle \geqq 0\ \forall x, y \in \triangle$, and that $\forall \xi \in \Lambda$, $sol(VI)_\xi \neq \emptyset$. If $sol(VVI)$ is bounded, then $sol(VVI)$ and $sol(VVI)^w$ are connected.*

Proof. Let $G : \Lambda \to 2^{I\!R^2}$ be a multifunction defined by $\forall \xi \in \Lambda$, $G(\xi) = \mathrm{sol(VI)}_\xi$. Then it follows from Proposition 1 that $G(\overset{o}{\Lambda}) = \mathrm{sol(VVI)}$ and $G(\Lambda) = \mathrm{sol(VVI)}^w$. Since $\sum_{i=1}^2 \xi_i M_i$ are monotone on $\triangle$, by Minty lemma ([6]), $G(\xi)$ is connected. Let $\bar{\xi} = (1,0)$. Suppose that G is not u.s.c. at $\bar{\xi}$. Then we can find an open subset Ω of $I\!R^2$, a sequence $\{\xi^k\}$ in $\overset{o}{\Lambda}$ and a sequence $\{x^k\}$ in $\triangle$ such that $G(\bar{\xi}) \subset \Omega$, $\xi^k \to \bar{\xi}$, $x^k \in G(\xi^k)$ and $x^k \notin \Omega$. Since $G(\xi^k) \subset \mathrm{sol(VVI)}$ and sol(VVI) is bounded, the sequence $\{x^k\}$ is bounded. So, without loss of generality, we may assume that $x^k \to \bar{x}$ for some $\bar{x} \in \triangle$. Since G is a closed multifunction, $\bar{x} \in G(\bar{\xi}) \subset \Omega$, However, since $x^k \notin \Omega$ for all k and Ω is open, $\bar{x} \notin \Omega$. This is a contradiction. Thus G is u.s.c. at $\bar{\xi}$. By the same (above) argument, we can check that G is u.s.c. at $\xi \in \Lambda \setminus \{(1,0)\}$. Hence G is u.s.c. on Λ. So by Lemma 1, sol(VVI) and $\mathrm{sol(VVI)}^w$ are connected.

Now we give examples to clarify our main result.

Example 1. Let

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad q_1 = q_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and $\triangle = \{(x_1, x_2) \in I\!R^2 \mid x_1 \geqq 0, x_2 \geqq 0\}$.

We consider (VVI), $\mathrm{(VVI)}^w$ and $\mathrm{(VI)}_{(\xi_1,\xi_2)}$ for the above M_i, $i = 1,2$, q_i, $i = 1,2$, and $\triangle$. Then $\forall (\xi_1, \xi_2) \in \Lambda, \mathrm{sol(VI)}_{(\xi_1,\xi_2)} \neq \emptyset$ and $\mathrm{sol(VVI)} = \{(0,0)\}$. So by Theorem 2.1, $\mathrm{sol(VVI)}^w$ is connected. Actually,
$\mathrm{sol(VVI)}^w = \{(x_1, 0) \mid x_1 \geqq 0\} \cup \{(0, x_2) \mid x_2 \geqq 0\}$.

Example 2. This example illustrates that the monotonicity assumption in Theorem 2.1 is essential. This example is slightly modified from the one of Robinson ([7]).
Let

$$M_1 = M_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad q_1 = q_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

and $\triangle = \{(x_1, x_2) \in I\!R^2 \mid x_1 - 2x_2 \geqq 0, x_1 + 2x_2 \geqq 0\}$.

Then M_1 and M_2 are not monotone on $\triangle$, $\forall (\xi_1, \xi_2) \in \Lambda, \mathrm{sol(VI)}_{(\xi_1,\xi_2)} = \{(1,0), (\frac{4}{3}, \frac{2}{3}), (\frac{4}{3}, -\frac{2}{3})\}$, and $\mathrm{sol(VVI)} = \mathrm{sol(VVI)}^w = \{(1,0), (\frac{4}{3}, \frac{2}{3}), (\frac{4}{3}, -\frac{2}{3})\}$. Thus sol(VVI) is bounded, but sol(VVI) and $\mathrm{sol(VVI)}^w$ are not connected.

Example 3. This example shows that the boundedness of sol(VVI) is essential in Theorem 2.1. This example is modified (came from) from the one of Choo and Atkins ([2]).
Let

$$M_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad q_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, q_2 = \begin{pmatrix} 0 \\ -3 \end{pmatrix},$$

and $\triangle = \{(x_1, x_2) \in I\!R^2 \mid x_1 \geqq 2, 0 \leqq x_2 \leqq 4\}$.

Then M_1 and M_2 are monotone on $\triangle$ and

$$\text{sol(VI)}_{(\xi_1,\xi_2)} = \begin{cases} \{(2,4)\} & \text{if } \ 0 \leqq \xi_1 < \frac{1}{2} \\ \{(x_1,4) \mid 2 \leqq x_1\}, & \text{if } \ \xi_1 = \frac{1}{2} \\ \{(x_1,0) \mid x_1 \geqq \frac{3-4\xi_1}{2\xi_1-1}\} & \text{if } \ \frac{1}{2} < \xi_1 < \frac{5}{8} \\ \{(x_1,0) \mid 2 \leqq x_1\} & \text{if } \ \frac{5}{8} \leqq \xi_1 \leqq 1. \end{cases}$$

By Proposition 1, we have

$$\begin{aligned} \text{sol(VVI)} &= \bigcup_{(\xi_1,\xi_2)\in \mathring{\Lambda}} \text{sol(VI)}_{(\xi_1,\xi_2)} \\ &= \bigcup_{(\xi_1,\xi_2)\in \Lambda} \text{sol(VI)}_{(\xi_1,\xi_2)} \\ &= \text{sol(VVI)}^w \\ &= \{(x_1,0) \mid 2 \leqq x_1\} \cup \{(x_1,4) \mid 2 \leqq x_1\}. \end{aligned}$$

So, sol(VVI) is not bounded, and sol(VVI) and sol(VVI)^w are not connected. It is worth while noticing that the multifunction $(\xi_1, \xi_2) \in \Lambda \to \text{sol(VI)}_{(\xi_1,\xi_2)}$ is not u.s.c. at $\xi = \frac{1}{2}$.

The converse of Theorem 1 may not be true.

Example 4. Let

$$M_1 = M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \quad q_1 = q_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and $\triangle = \{(x_1, x_2) \in I\!R^2 \mid x_1 \geqq 0, x_2 \geqq 0\}$.

Then M_1 and M_2 are monotone on $\triangle$, $\forall(\xi_1, \xi_2) \in \Lambda$, $\text{sol(VI)}_{(\xi_1,\xi_2)} = \{(0, x_2) \in I\!R^2 \mid x_2 \geqq 0\}$, and $\text{sol(VVI)} = \text{sol(VVI)}^w = \{(0, x_2) \in I\!R^2 \mid x_2 \geqq 0\}$. Thus sol(VVI) is not bounded, but sol(VVI) and sol(VVI)^w are connected.

QUESTION: Can we find an example which shows that Theorem 2.1 does not hold for affine vector variational inequalities for monotone 3×3 matrices?

References

1. Cheng, Y. (2001) On the connectedness of the solution set for the weak vector variational inequality, J. Math. Anal. Appl. **260**, 1–5.

2. Choo, E. U. and Atkins, D. R. (1983) Connectedness in multiple linear fractional programming, Management Science **29**, 250–255.
3. Giannessi, F. (1980) Theorems of alternative, quadratic programs and complementarity problems, in "Variational Inequalities and Complementarity Problems" Edited by R.W. Cottle, F. Giannessi, and J. L. Lions, Wiley, Chichester, England, pp. 151-186.
4. Lee, G. M., Kim, D. S., Lee B. S. and Yen, N. D. (1998) Vector variational inequality as a tool for studying vector optimization problems, Nonlinear Analysis **34**, 745–765.
5. Lee, G. M. and Yen, N. D. (2001) A result on vector variational inequalities with polyhedral constraint sets, J. Optimz. Th. Appl. **109**, 193–197.
6. Minty, G. J. (1962) Monotone (nonlinear) operators in Hilbert space, Duke Math.J. **29**, 341–346.
7. Robinson, S. M. (1980) Strongly regular generalized equations, Math. Oper. Res. **5**, 43–62.
8. Warburton, A. R. (1983) Quasiconcave vector maximization: Connectedness of the sets of Pareto-optimal and weak Pareto-optimal alternatives, J. Optimiz. Th. Appl. **40**, 537–557.
9. Yen, N. D. and Phuong, T. D. (2000) Connectedness and stability of the solution sets in linear fractional vector optimization problems, in "Vector Variational Inequalities and Vector Equilibria", edited by F. Giannessi, pp. 479–489, Kluwer Academic Publishers, Dordrecht/ Boston/ London.
10. Yen, N. D. and Lee, G. M. On monotone and strongly monotone vector variational inequalities, pp. 467-478, in the above book.
11. Yu, M., Wang, S. Y., Fu, W. T. and Xiao, W. S. (2001) On the existence and the connectedness of solution sets of vector variational inequalities, Math. Meth. Oper. Res. **54**, 201-215.

Graphical Illustration of Pareto Optimal Solutions

Kaisa Miettinen

University of Jyväskylä, Department of Mathematical Information Technology
P.O. Box 35 (Agora), FIN-40014 University of Jyväskylä, Finland

Abstract. The comparison of different alternatives is a part of many multicriteria optimization and decision making methods. This task may be very demanding for the decision maker and illustrating the alternatives is often helpful. Unfortunately, with more than two criteria, the graphical illustration is not trivial and it gets even more demanding with more alternatives and/or more criteria. Here, we summarize several ways of illustrating a set of alternatives graphically.

1 Introduction

Real-life optimization problems often involve several conflicting criteria that should be minimized or maximized simultaneously. In such multicriteria optimization problems, there usually exist many mathematically equivalent, so-called Pareto optimal or efficient solutions, and a human decision maker (DM) is required in order to find the final solution among them. This means that the most preferred solution is identified based on the preferences of the DM.

Quite a few interactive multicriteria methods (see, e.g., [10] and references therein) assume the DM to select the most preferred alternative from a given set as a part of the solution process but they do not necessarily assist the DM in this comparison task. However, usually, the comparison problem is difficult to be solved directly. In these cases, we need tools that describe different features and elements involved in a simple but rigorous way and, here, graphical illustration of the alternatives is a noteworthy tool. It can be used in exploring the data in order to gain insight into the data itself as well as understanding of the underlying phenomena and the problem solving process.

In this paper, we treat graphical illustration of alternative solutions with the goal of supporting comparison. In the literature, one can find surprisingly seldom methods where graphical illustration is used in assisting the DM. Short summaries of graphical illustration tools are given, for example, in [6,7,13] but they mostly concentrate on one or two tools and only mention some of the others. That is why it is in order to present a general summary of the possibilities available. The aim here is to help all those who need an overview of the existing possibilities including those who develop multicriteria methods. The potential and restrictions of different graphics tools are treated and some clarifying figures are enclosed. The presentation is based on the background laid in [10].

In what follows, we assume that we have at least two criteria and a finite set of alternatives that we wish to illustrate. The alternatives consist of criterion values. We assume that lower criterion values are preferred to higher in each criterion.

In many illustrations, we need to know the ranges of the criterion values in the Pareto optimal set. A vector consisting of the best (that is, smallest) values of each criterion is called an *ideal criterion vector*. Correspondingly, the worst values are the components of a *nadir criterion vector*. How these vectors are formed, depends on the form of the problem and we do not touch that topic here. Note that in many occasions, it is advisable to normalize the criteria. With the help of the ideal and the nadir criterion vectors, it is easy to normalize the criteria so that their ranges equal $[0, 1]$.

2 Graphical Illustration

The ultimate goal of graphical illustration is to enable the DM to gain more understanding of the problem and new insight into the alternatives and the underlying phenomena. It is essential that the graphics must be clear as well as easy to comprehend and interpret by the DM. On the one hand, not too much information should be allowed to be lost and, on the other hand, no extra unintentional information should be included in the presentation.

Our intention is to show several possibilities for illustrating alternatives so that the DM can more easily differentiate between them. The graphical illustrations can bring out similarities and differences, which helps the DM in dropping uninteresting alternatives and identifying the most preferred one.

Nevertheless, utilizing graphical illustration does not mean that the limits on human information processing capacity are transcended. Several psychological tests are summarized in [12] to prove that the span of absolute judgement and the span of immediate memory in human beings is rather limited. We cannot receive, process or remember large amounts of information. As stressed in [8], experiments in psychology indicate that the amount of information provided to the DM has a crucial role. If more information is given to the DM, 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 DM in the solution obtained but the quality of the solution may nonetheless be worse.

In what follows, we briefly discuss several possibilities of illustrating graphically a given set of alternative criterion vectors. We also give examples of these graphical tools. Unfortunately, because of lack of space, we can only show the general appearance of each tool for one alternative. Naturally, this does not give the right impression about the general usability of the tools in comparison.

Value Path In *value paths*, [4,13] horizontal lines of different colours or of different line styles represent the values of the criteria in different alternatives. In other words, one line is associated with one alternative. This is depicted in Fig. 1(a). The bars in the figure show the ranges of the criteria in the Pareto optimal set. Thus, they give additional information about the goodness of the criterion values at hand. Here, each criterion can have a scale of its own in the bars, if necessary. Note that the roles of the lines and the bars can also be interchanged so that bars denote alternatives and lines represent criteria.

Value paths are used, for example, in [2] and in the interactive multiobjective optimization system WWW-NIMBUS http://nimbus.mit.jyu.fi/ [11].

Bar Chart In *bar charts*, a group of bars represents the alternative values of a single criterion. The bars of the same colour are related to one alternative. Separate ranges for criteria are possible as well. Naturally, the roles of the alternatives and the criteria can be interchanged so that the bars are grouped according to alternatives instead of criteria (as in Fig. 1(b)).

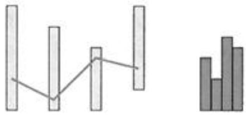


Fig. 1. Value path (**a**) and bar chart (**b**)

Bar charts take horizontally quite a lot of space. In this case, the bars may be located, for example, in three dimensions [11] or in a vertical direction. The flexibility is increased if the DM can affect the order of the alternatives, that is, the order and the assignment of the bars.

Star Coordinate System In *star coordinate system* [9] or *star presentations* [14], rays emanating from the origin represent criteria. For example, an alternative involving three criteria is represented as an irregular triangle. An example with four criteria is given in Fig. 2(a). Each circle represents one alternative and the area of each star depicts the goodness of that alternative. In each ray, the ideal criterion value is located at the centre and the nadir criterion value is at the circumference.

If the areas are not filled, we can locate several alternatives in the same circle. Up to some point, this may make the comparison easier (it is evident that too many alternatives cannot fit in one circle). Note that if the order of the criteria is altered, the shape and the area of the star change. This can be considered a weakness of the system [15]. Alternatively, it is possible to only display the line segments along the rays, as in Fig. 2(b). Naturally, the order of the criteria affects the appearance even then but the strong association with areas is avoided.

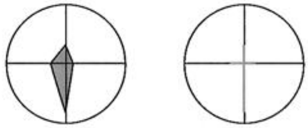


Fig. 2. Star coordinate system (**a**) and the same with line segments (**b**)

Spider-Web Chart In *spider-web* or *radar charts* [5], each apex of a polygon represents one criterion. For example, with three criteria we have triangles in question. An example is presented in Fig. 3(a). The outer polygon shows the nadir criterion vector, the inner polygon stands for the ideal criterion vector and the middle polygon (the grey one) represents one alternative criterion vector. Thus, only the middle polygon is different in each chart. Note that one can also locate the polygons of several alternatives in the same chart.

Petal Diagram In *petal diagrams* [15], a circle is divided into as many equal sectors as there are criteria. The size (radius) of each slice indicates the magnitude of the criterion value. Thus, we have one circle for each alternative. Each segment of the diagram, that is, each criterion can be associated with a different colour, as in Fig. 3(b). Notice that the order of the criteria has no effect on the actual shape of the diagram or the total area covered by the segments. Petal diagrams are utilized, for example, in [1,3,11].

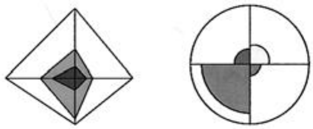


Fig. 3. Spider-web chart (**a**) and petal diagram (**b**)

It is mainly a matter of taste in the star coordinate system, the spider-web chart and the petal diagram, how the ideal criterion vector is situated. The roles can be interchanged so that the ideal criterion value is located on the circumference and the nadir criterion value at the centre. In this case, the larger the area the better. Alternatively, one can locate some reference values (if available) on the circumference. In this case, the figures may extend beyond the circumference and, thus, the DM can easily identify desirable criterion values.

Because of lack of space, we do not here touch projection or icon based approaches or other illustrative means.

3 Discussion

There is no straight-forward answer to the question when to use which type of graphical illustrations. The choice is up to the DMs and the problems in question. Even though new graphical tools may bring along new possibilities, they necessitate time for training the DMs in interpreting them.

One should not forget that tables of alternative data may still be needed because tables usually perform better in information acquisition tasks whereas graphs are valuable in viewing data at a glance or evaluating relationships in the data. Thus, one can say that graphs and tables emphasize different characteristics of the same data and they complement each other.

People often prefer colourful pictures but it is not always clear that the colours will make the pictures easier to comprehend. Above all, the colours must be easy to discriminate. The advantage of colours is that they make it easier for the DM to visually associate information belonging to the same context like alternative or criterion. Unfortunately, it is very easy to overload the DM with too much colour information.

A recommended way of presenting information to the DM is to offer the same data in different forms. A simple tabular format may be one of the figures. In this way, the DM can choose the most illustrative and informative representations. The illustrations may also supplement each other. This idea is used WWW-NIMBUS [11].

4 Conclusions

Many solution approaches for multicriteria optimization problems involve comparison of alternatives, that is, criterion vectors. Yet, the DM is not necessarily supported in this task. One possibility to help the DM in selecting the most preferred alternative is to use graphical illustration of the alternatives. For some reason, illustration is used relatively seldom. Some explanation may be found in the fact that there are no surveys in the literature of the appropriate illustrative tools available. That is why we have here presented a collection of different possibilities for graphical illustration. The idea has been to provide a summary as a starting point for those willing to illustrate alternatives graphically. None of the graphical representations can be claimed to be better than the others but some fit certain problem types better than the others. It is always good to leave the final decision to the DM who can select those illustrations that (s)he is most comfortable with.

Acknowledgements This research was supported by the grant 65760 of the Academy of Finland.

References

1. Angehrn, A.A. (1991): Designing Humanized Systems for Multiple Criteria Decision Making. Human Systems Management 10, 221–231
2. Belton, V., Vickers, S. (1990): Use of a Simple Multi-Attribute Value Function Incorporating Visual Interactive Sensitivity Analysis for Multiple Criteria Decision Making. In: Bana e Costa, C.A. (Ed.): Readings in Multiple Criteria Decision Aid. Springer, Berlin Heidelberg, 319–334
3. Gandibleux, X. (1999): Interactive Multicriteria Procedure Exploiting a Knowledge-Based Module to Select Electricity Production Alternatives: The CASTART System. European Journal of Operational Research 113, 355–373
4. Geoffrion, A.M., Dyer, J.S., Feinberg, A. (1972): An Interactive Approach for Multi-Criterion Optimization, with an Application to the Operation of an Academic Department. Management Science 19, 357–368
5. Kasanen, E., Östermark, R., Zeleny, M. (1991): Gestalt System of Holistic Graphics: New Management Support View of MCDM. Computers & Operations Research 18, 233–239
6. Klimberg, R. (1992): GRADS: A New Graphical Display System for Visualizing Multiple Criteria Solutions. Computers & Operations Research 19, 707–711
7. Klimberg, R., Cohen, R.M. (1999): Experimental Evaluation of a Graphical Display System to Visualizing Multiple Criteria Solutions. European Journal of Operational Research 119, 191–208
8. Kok, M. (1986): The Interface with Decision Makers and Some Experimental Results in Interactive Multiple Objective Programming Method. European Journal of Operational Research 26, 96–107
9. Mañas, M. (1982): Graphical Methods of Multicriterial Optimization. Zeitschrift für Angewandte Mathematik und Mechanik 62, 375–377
10. Miettinen, K. (1999): Nonlinear Multiobjective Optimization. Kluwer, Boston
11. Miettinen, K., Mäkelä, M.M. (2000): Interactive Multiobjective Optimization System WWW-NIMBUS on the Internet. Computers & Operations Research 27, 709–723
12. Miller, G.A. (1956): The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. Psychological Review 63, 81–87
13. Schilling, D.A., ReVelle, C., Cohon, J. (1983): An Approach to the Display and Analysis of Multiobjective Problems. Socio-Economic Planning Sciences 17, 57–63
14. Sobol, M.G., Klein, G. (1989): New Graphics as Computarized Displays for Human Information Processing. IEEE Transactions on Systems, Man, and Cybernetics 19, 893–898
15. Tan, Y.S., Fraser, N.M. (1998): 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, 20–33

An Efficiency Evaluation Model for Company System Organization

Takashi Namatame[1], Hiroaki Tanimoto[2], and Toshikazu Yamaguchi[3]

[1] Senshu University, Kanagawa 214-8580, JAPAN
[2] Hitachi, Ltd., Kanagawa 211-8567, JAPAN
[3] Tokyo University of Secince, Tokyo 162-8601, JAPAN

Abstract. The company system organization is focused as a decentralization organization of the company recently. Generally, the company system organization has headquarters, companies and divisions from the top. So, It has multiple hierarchical structure. Hence, an integrated evaluation for the company system organization is not simple. In this paper, we propose an efficiency evaluation model for company system organization in consideration of the characteristics of the company system organization via data envelopment analysis concept.

1 Introduction

Many Japanese enterprises have been doing the rationalizations of the structure of the various management aspects after the bubble economic collapse. The company system organization is focused as a trend which the structure of the management is changed into from the side of the reform of the organization [2]. In the company sysmtem organization, the decentralization of the business management is done, and it has responsibility for a business cleared, so each company aims to increase its management efficiency. It is useful to know what kind of activities each company and division under the headquaters is doing in the enterprise.

In this paper, we evaluate relatively the activity of company system organization via data envelopment analysis (DEA) [1] which evaluate the relative efficiency among plural decision making units (DMUs). Though various DEA models have been proposed, there is no DEA model for company system organization. So, we propose an efficiency evaluation model for company system organization. We utilize the cross-efficiency to evaluate the companies, integratively. We propose an improved method to obtain the unique weights for the cross-efficiency.

2 Characteristics of the Company System Organization

We enumerate the characteristics of the company system organization in this paper.

1. The company has three hierarchical structure about management, i.e. headquaters, company and division.

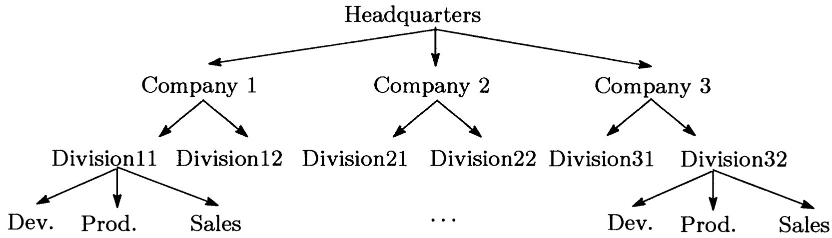


Fig. 1. Outline of company system organization

2. Each company is the business unit that has the authority which is close to an independent firm, and responsibility.
3. All divisions have the same sections (functions), and the inputs and outputs of division are distributed into the sections.
4. Same section has same inputs and outputs.

Figure 1 shows the outline of company system organization. Headquarters plans the whole enterprise strategies, and entrusts the measure and practice of executive strategies to each company. Each company entrusts to practice the business strategies and managment to divisions under the company. Each division practices business strategies by controlling its sections. Each company is managed just like an independent firm in the enterprise. So, each company want to evaluate fairly among all companies, but ones best as much as possible. Therefore, when the companies are evaluated, they are expected to hold the conditions that each other can cut in about the evaluation criterion, so an evaluation criterion may be difficult to be understood by the other company.

3 Evaluation Model

We propose an integrated efficiency evaluation model in consideration of the above characteristics and structure. The evaluation process is roughly divided into three steps. In step 1, we evaluate each division. In step 2, we evaluate each company by the integrated criterion based on the mutual evaluation information among companies. In step 3, we evaluate each company from the headquarters' point of view.

3.1 Variables

ℓ_j, o and a are subscripts that mean the division in the jth company, target company and the target division, n and s_j are the number of companies and divisions in the company j, respectively. q is the number of sections. m and k are the number of inputs and outputs, m_p and k_p are the number of inputs and outputs of the pth section. So, the sum of m_p and k_p with respect to p are m and k, respectively. $X_{i\ell_j j}$, $Y_{r\ell_j j}$ are the ith input and the rth output of the ℓ_jth division in the jth company, respectively. $v_{i\ell_j j}$ and $u_{r\ell_j j}$ are the decision variables which weights for the ith input and the rth output of the ℓ_jth division in the jth company.

3.2 Evaluation Procedure

We explain the detail of the evaluation procedure in the following.

Step 1. We evaluate the efficiency of each division on the whole by [P1] in accordance with Yang et al. [5] which is a DEA model with the subsystems.

$$[\text{P1}] \quad (a = 1, \ldots, s_o; o = 1, \ldots, n)$$

$$\max \ h_{ao} = \frac{\sum_{r=1}^{k} u_{rao} Y_{rao}}{\sum_{i=1}^{m} v_{iao} X_{iao}} \tag{1a}$$

$$\text{s.t.} \ \frac{\sum_{r=k_{p-1}+1}^{k_p} u_{rao} Y_{r\ell_j j}}{\sum_{i=m_{p-1}+1}^{m_p} v_{iao} X_{i\ell_j j}} \leq 1,$$

$$j = 1, \ldots, n; p = 1, \ldots, q; \ell_j = 1, \ldots, s_j \tag{1b}$$

$$v_{iao}, u_{rao} \geq 0, \quad i = 1, \ldots, m; r = 1, \ldots, k \tag{1c}$$

This formulation maximizes the efficiency score of the target division under the constraints that the efficiency scores of all sections (subsystems) is less than or equal unity.

Step 2. The inputs and outputs of all divisions under each company are collected, and we calculate the efficiency score of each company. The inputs and outputs of company j, X_{ij} and Y_{rj}, are the sum of input and output of all divisions under each compnay, respectively, that is $X_{ij} = \sum_{\ell_j=1}^{s_j} X_{i\ell_j j}$ and $Y_{rj} = \sum_{\ell_j=1}^{s_j} Y_{r\ell_j j}$ with respct to $i = 1, \ldots, m$, $r = 1, \ldots, k$ and $j = 1, \ldots, n$.

In this step, to integrate the evaluation criterion, the accommodation of the evaluation of each company is done by using the mutual evaluation information among the companies.

The cross-efficiency [2] is a mutual evaluation method which the mutual evaluation information among DMUs is used for. But, the cross-efficiency has a serious problem when the optimal weights are not unique, its score is not determined uniquely . To this problem, Hibiki [3] has proposed a way to determine the optimal weights. His method is need a decision maker's judgement for score in advance. But, it is not so easy to get the proper value. So, we propose a method which obtains the cross-efficiency score uniquely without decision maker's judgement based on his method thruough the next sub-steps.

Step 2-1. The efficiency score of each company θ_o^* is calculated by the ordinal DEA model.

Step 2-2. If the optimal weight is unique then the cross-efficiency value is obtained by equation (2), else go to step 2-3.

$$E^*_{jo} = \frac{\sum_{r=1}^{k} u^*_{ro} Y_{rj}}{\sum_{i=1}^{m} v^*_{io} X_{ij}}, \quad j = 1, \ldots, n; o = 1, \ldots, n \tag{2}$$

where v^*_{io}, u^*_{ro} are the optimal weights of the ith input and the rth output of DMU_o, respectively.

Step 2-3. We obtain the maximum and minimum cross-efficincy score of the other DMUs with holding θ^*_o by solving [P2] and [P3], respectively.

[P2] $(o = 1, \ldots, n; b = 1, \ldots, n; b \neq o)$

$$\max \ E^U_{bo} = \frac{\sum_{r=1}^{k} u_{rb} Y_{rb}}{\sum_{i=1}^{m} v_{ib} X_{ib}} \tag{3a}$$

$$\text{s.t.} \ \frac{\sum_{r=1}^{k} u_{rb} Y_{ro}}{\sum_{i=1}^{m} v_{ib} X_{io}} = \theta^*_o \tag{3b}$$

$$\frac{\sum_{r=1}^{k} u_{rb} Y_{rj}}{\sum_{i=1}^{m} v_{ib} X_{ij}} \leq 1, \quad j = 1, \ldots, n; j \neq o \tag{3c}$$

$$v_{ib}, u_{rb} \geq 0, \quad i = 1, \ldots, m; r = 1, \ldots, k \tag{3d}$$

[P3] $(o = 1, \ldots, n; b = 1, \ldots, n; b \neq o)$

$$\min \ E^L_{bo} = \frac{\sum_{r=1}^{k} u_{rb} Y_{rb}}{\sum_{i=1}^{m} v_{ib} X_{ib}} \tag{4}$$

s.t. Equations (3b)–(3d)

Step 2-4. The optimal weights for the cross-efficiency is obtained by [P4].

[P4] $(o = 1, \ldots, n)$

$$\max \ d_o \tag{5a}$$

$$\text{s.t.} \ \frac{\sum_{r=1}^{k} u_{ro} Y_{ro}}{\sum_{i=1}^{m} v_{io} X_{io}} = \theta^*_o \tag{5b}$$

$$\frac{\sum_{r=1}^{k} u_{ro} Y_{rj}}{\sum_{i=1}^{m} v_{io} X_{ij}} \geq E^{L*}_{jo} + \left(E^{U*}_{jo} - E^{L*}_{jo}\right) d_o, \quad j = 1, \ldots, n; j \neq o \tag{5c}$$

$$v_{io}, u_{ro} \geq 0, \quad i = 1, \ldots, m; r = 1, \ldots, k \tag{5d}$$

where E^{U*}_{jo} and E^{L*}_{jo} are the optimal value of [P2] and [P3]. [P4] can not be transformed to a linear programming problem, but optimal solution of [P4] is in [0, 1], so can be solved by using simplex method and bisection method. Then, optimal cross-efficinecy score E^*_{jo} is obtained by equation (2) with respect to the optimal solution of [P4].

The average of the cross-efficiency score of each DMU is finally made the efficiency of each company as equation (6).

Table 1. Example data

Div.	I1	I2	O	Div.	I1	I2	O
Company 1				Company 3			
	Development section				Development section		
L	320	65	250	Q	700	140	200
M	183	7	200	R	450	110	200
	Product section				Product section		
L	350	150	250	Q	750	126	330
M	200	15	150	R	650	123	180
	Sales section				Sales section		
L	330	335	175	Q	717	250	254
M	200	210	33	R	400	250	227
Company 2				Company 4			
	Development section				Development section		
N	800	90	300	S	650	80	150
O	300	85	300	T	400	90	200
P	210	67	250		Product section		
	Product section			S	650	90	100
N	410	57	120	T	600	70	150
O	200	53	250		Sales section		
P	130	34	110	S	700	170	80
	Sales section			T	500	200	50
N	456	90	163				
O	166	130	112				
P	160	90	59				

$$\bar{E}_j = \frac{1}{n}\sum_{o=1}^{n} E^*_{jo}, \quad j = 1, \dots, n \tag{6}$$

Step 3. We evaluate each company from the headquarters' point of view, synthetically, based on the results of Step 1 and 2.

4 Example

We show a simple exapmle in table 1. The enterprise has four companies, company 1, 2, 3 and 4. Company 1, 3 and 4 has each two divisions, and company 2 has three one. Each division has three sections, that is development, product and sales. Each section has two inputs and one output. Input 1 (I1), 2 (I2) are common for all sections, the employee and budget. Output (O) is the number of patents, the number of product, and the sales ammount for development, product and sales section, respectively.

From step 1, the efficiency score of division L, Q, R, S and T are 0.7860, 0.8602, 0.9634, 0.5991 and 0.4319, respectively, and the others are unity.

When we applying CCR model, only divisions S and T are inefficient that the scores are 0.5305 and 0.5968, respectively. From this result, the number of inefficient DMUs increase.

The weights of company 1, 2 and 3 is not uniquely. The next is the cross-efficinecy matrix by our method.

$$[E^*_{jo}] = \begin{bmatrix} 1 & 0.9997 & 0.9999 & 1 \\ 0.9995 & 1 & 0.9999 & 1 \\ 0.9396 & 0.6423 & 1 & 0.8529 \\ 0.6025 & 0.3014 & 0.5823 & 0.6325 \end{bmatrix}$$

The final company scores are $\bar{E}_j$ are 0.9999, 0.9998, 0.9337 and 0.6047, respectively.

From the results for evaluation of company and division, we can evaluate company 2 to be the most efficient, synthetically. Compnay 3 is evaluated almost efficient by our method, but the divisions under company 3 are inefficient. So, we need some attention.

5 Conclusion

In this paper, we have proposed an evaluation model for company system organization. In our method, we have considered the multiply hierarchy structure of company system organization and we proposed a way to determine the optimal weight for the cross-efficiency.

We treat static case, in this paper. The activity of enterprise always moves, so we should extend our method to time-series evaluation.

References

1. Cooper, W.W., Seiford, L.M., Tone, K. (1999) Data Envelopment Analysis—A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Kluwer Academic Publishers, Dordrecht
2. Fushimi, T., Watanabe, Y. (1995) A note on "Kanpani-sei" management: survey from the viewpoint of management control systems. Keio Business Forum, **13**(1), 45–74 (in Japanese)
3. Hibiki, N. (1998) Evaluation Techniques with a Modified Cross-Efficiency in DEA. J. of the Oper. Res. Soc. of Japan. **41**, 229–245 (in Japanese)
4. Sexton, T.R., R.H. Silkman, A.J. Hogan (1986) Data envelopment analysis: critique and extensions. (Silkman T.H., Ed.) Measureing Efficiency: Assessment of DEA, Jossy Bass, San Francisco. 73–105
5. Yang, Y., Ma, B., Koike, M. (2000) Efficiency-measuring DEA model for production system with k independet subsystem. J. of the Oper. Res. Soc. of Japan. **43**, 343–354

Stackelberg Solutions to Two-Level Linear Programming Problems with Random Variable Coefficients

Ichiro Nishizaki, Masatoshi Sakawa, Kosuke Kato, and Hideki Katagiri

Department of Artificial Complex Systems Engineering, Graduate School of Engineering, Hiroshima University, Higashi-Hiroshima, 739-8527, Japan

Abstract. In this paper, to cope with hierarchical decision making problems under uncertainty, we formulate a two-level linear programming problem in which random variable coefficients are involved in objective functions and constraints, and reduce the problem into deterministic problems by using two models. While one of the deterministic problems is a usual two-level linear programming problem, the other is a two-level quadratic one. We present a computational method for obtaining Stackelberg solutions to the reduced deterministic two-level quadratic programming problems.

1 Introduction

In real-world decision making problems, there are many uncertain elements and coefficients of the formulated mathematical models cannot be always determined precisely. From the viewpoint, we deal with two-level linear programming problems with random variable coefficients in this paper. Charnes and Cooper [5] proposed stochastic programming models from various different viewpoints: E-model, V-model and P-model. The E-model aims at optimizing the expected value of an objective function and the V-model minimizes the variance of an objective function value. The P-model maximizes the probability that an objective function is smaller than an aspiration level in the minimization problem.

From the point of view that we give the leader some advice on how to make a decision, we reduce two-level linear programming problems with random variable coefficients into certain deterministic problems by using the E- and the V-models. Because the deterministic problem corresponding to the V-model is a two-level quadratic programming problems, we develop a computational method for obtaining Stackelberg solutions to the reduced deterministic two-level quadratic programming problems.

2 Two-level linear programming problems with random variable coefficients

Let $\boldsymbol{x} \in \mathbb{R}^{n_1}$ and $\boldsymbol{y} \in \mathbb{R}^{n_2}$ denote a pair of decision variable column vectors of the leader and the follower. We deal with the following two-level linear programming problems with random variable coefficients:

$$\left.\begin{array}{l} \underset{\boldsymbol{x}}{\text{minimize}}\ z_1(\boldsymbol{x},\boldsymbol{y}) = \tilde{\boldsymbol{c}}_1\boldsymbol{x} + \tilde{\boldsymbol{d}}_1\boldsymbol{y} \\ \text{where } \boldsymbol{y} \text{ solves} \\ \underset{\boldsymbol{y}}{\text{minimize}}\ z_2(\boldsymbol{x},\boldsymbol{y}) = \tilde{\boldsymbol{c}}_2\boldsymbol{x} + \tilde{\boldsymbol{d}}_2\boldsymbol{y} \\ \text{subject to } A_1\boldsymbol{x} + A_2\boldsymbol{y} \leq \tilde{\boldsymbol{b}} \\ \qquad \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0}, \end{array}\right\} \tag{1}$$

where $z_1(\boldsymbol{x},\boldsymbol{y})$ is an objective function of the leader; $z_2(\boldsymbol{x},\boldsymbol{y})$ is an objective functions of the follower; $\tilde{\boldsymbol{c}}_1$, $\tilde{\boldsymbol{c}}_2$ are n_1-dimensional row vectors of random variable coefficients; $\tilde{\boldsymbol{d}}_1$, $\tilde{\boldsymbol{d}}_2$ are n_2-dimensional row vectors of random variable coefficients; A_1 and A_2 are $m \times n_1$ and $m \times n_2$ coefficient matrices; $\tilde{\boldsymbol{b}}$ is an m-dimensional column vector of random variable coefficients.

Because problem (1) contains random variable coefficients, definitions and solution methods for ordinary mathematical programming problems cannot be directly applied. In this paper, for the constraints, we employ the concept of the chance constrained conditions [4]. Namely, the probability that the constraints are satisfied is not less than a given probability level. Let α_i be a probability level for the ith constraint specified by the leader. The chance constrained condition is represented by

$$\Pr\{A_1^i\boldsymbol{x} + A_2^i\boldsymbol{y} \leq \tilde{b}_i\} \geq \alpha_i,\ i = 1,\ldots,m, \tag{2}$$

where A_1^i, A_2^i and $\tilde{b}_i$ are coefficients of the ith constraint. Let $F_i(\tau)$ be a distribution function of the random variable $\tilde{b}_i$. Since $\Pr\{A_1^i\boldsymbol{x} + A_2^i\boldsymbol{y} \leq \tilde{b}_i\} = 1 - F(A_1^i\boldsymbol{x} + A_2^i\boldsymbol{y})$, the inequality (2) is rewritten as $F(A_1^i\boldsymbol{x} + A_2^i\boldsymbol{y}) \leq 1 - \alpha_i$. Let $K_{1-\alpha_i}$ denote the maximal value of τ satisfying $\tau = F^{-1}(1-\alpha_i)$. From monotonicity of the distribution function, the inequality (2) can be transformed into

$$A_1^i\boldsymbol{x} + A_2^i\boldsymbol{y} \leq K_{1-\alpha_i},\ i = 1,\ldots,m. \tag{3}$$

Then, we have the following problem:

$$\left.\begin{array}{l} \underset{\boldsymbol{x}}{\text{minimize}}\ z_1(\boldsymbol{x},\boldsymbol{y}) = \tilde{\boldsymbol{c}}_1\boldsymbol{x} + \tilde{\boldsymbol{d}}_1\boldsymbol{y} \\ \text{where } \boldsymbol{y} \text{ solves} \\ \underset{\boldsymbol{y}}{\text{minimize}}\ z_2(\boldsymbol{x},\boldsymbol{y}) = \tilde{\boldsymbol{c}}_2\boldsymbol{x} + \tilde{\boldsymbol{d}}_2\boldsymbol{y} \\ \text{subject to } A_1\boldsymbol{x} + A_2\boldsymbol{y} \leq \boldsymbol{K}_{1-\alpha} \\ \qquad \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0}, \end{array}\right\} \tag{4}$$

where $\boldsymbol{K}_{1-\alpha} = (K_{1-\alpha_1},\ldots,K_{1-\alpha_m})^T$ and the superscript T represents transposition of vectors or matrices.

For problem (4), we examine the E- and the V-models. In the E-model, the expected values of the objective functions $z_1(\boldsymbol{x},\boldsymbol{y})$ and $z_2(\boldsymbol{x},\boldsymbol{y})$ of the leader and the follower are minimized, and the corresponding deterministic problem is formulated as:

$$\left.\begin{array}{l} \underset{\boldsymbol{x}}{\text{minimize}}\ E[\tilde{\boldsymbol{c}}_1\boldsymbol{x} + \tilde{\boldsymbol{d}}_1\boldsymbol{y}] = \boldsymbol{m}_1^c\boldsymbol{x} + \boldsymbol{m}_1^d\boldsymbol{y} \\ \text{where } \boldsymbol{y} \text{ solves} \\ \underset{\boldsymbol{y}}{\text{minimize}}\ E[\tilde{\boldsymbol{c}}_2\boldsymbol{x} + \tilde{\boldsymbol{d}}_2\boldsymbol{y}] = \boldsymbol{m}_2^c\boldsymbol{x} + \boldsymbol{m}_2^d\boldsymbol{y} \\ \text{subject to } A_1\boldsymbol{x} + A_2\boldsymbol{y} \leq \boldsymbol{K}_{1-\alpha} \\ \qquad \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0} \end{array}\right\} \tag{5}$$

where $E[f]$ denotes the expected value of f; $\boldsymbol{m}_i^c$, $\boldsymbol{m}_i^d$, $i = 1,2$ are vectors of expected values of $\tilde{\boldsymbol{c}}_i$, $\tilde{\boldsymbol{d}}_i$, $i = 1,2$. Because problem (5) is a usual two-level linear programming problem, algorithms previously developed can be applied [1,3,2,6,7].

Although the expected value of the objective function value is minimized as in the E-model, it does not seem that the obtained solution is appropriate for the decision making under uncertainty if the variance is considerably large. In such a case, it is natural to employ the variance minimization model, i.e., the V-model. In the V-model, the variances of the objective functions $z_1(\boldsymbol{x},\boldsymbol{y})$ and $z_2(\boldsymbol{x},\boldsymbol{y})$ of the leader and the follower are minimized and it is often desirable to search solutions satisfying certain levels specified by the leader for the expected values of the objective functions. Such a problem can be represented by

$$\left.\begin{array}{l} \underset{\boldsymbol{x}}{\text{minimize}}\ Var[\tilde{\boldsymbol{c}}_1\boldsymbol{x} + \tilde{\boldsymbol{d}}_1\boldsymbol{y}] = \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}^T V_1 \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} \\ \text{where } \boldsymbol{y} \text{ solves} \\ \underset{\boldsymbol{y}}{\text{minimize}}\ Var[\tilde{\boldsymbol{c}}_2\boldsymbol{x} + \tilde{\boldsymbol{d}}_2\boldsymbol{y}] = \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}^T V_2 \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} \\ \text{subject to } A_1\boldsymbol{x} + A_2\boldsymbol{y} \leq \boldsymbol{K}_{1-\alpha} \\ \qquad \boldsymbol{m}_1^c\boldsymbol{x} + \boldsymbol{m}_1^d\boldsymbol{y} \leq \beta_1 \\ \qquad \boldsymbol{m}_2^c\boldsymbol{x} + \boldsymbol{m}_2^d\boldsymbol{y} \leq \beta_2 \\ \qquad \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0} \end{array}\right\} \tag{6}$$

where $Var[f]$ denotes the variance of f; V_i, $i = 1,2$ are covariance matrices with respect to $(\tilde{\boldsymbol{c}}_i, \tilde{\boldsymbol{d}}_i)$, $i = 1,2$; β_1 and β_2 are satisfactory level for the expected values of the objective function of the leader and the follower.

We consider a computational method for obtaining Stackelberg solutions in the V-model. The leader chooses a decision so as to minimize the variance of the objective function value on the assumption that, for the given decision of the leader, the follower takes an optimal response, i.e., a decision such that the decision minimizes the variance of the follower's objective function value. We develop a computational method for obtaining such a pair of decisions, i.e., the Stackelberg solution.

After the leader has chosen a decision $\hat{\boldsymbol{x}}$, if the follower intends to minimize the variance of the objective function, it follows that the follower chooses an optimal solution to the problem

$$\left.\begin{array}{l} \text{minimize } Var[\tilde{\boldsymbol{c}}_2\hat{\boldsymbol{x}} + \tilde{\boldsymbol{d}}_2\boldsymbol{y}] = \begin{bmatrix}\hat{\boldsymbol{x}}\\ \boldsymbol{y}\end{bmatrix}^T V_2 \begin{bmatrix}\hat{\boldsymbol{x}}\\ \boldsymbol{y}\end{bmatrix} \\ \text{subject to } A_1\hat{\boldsymbol{x}} + A_2\boldsymbol{y} \leq \boldsymbol{K}_{1-\alpha} \\ \qquad \boldsymbol{m}_1^c\hat{\boldsymbol{x}} + \boldsymbol{m}_1^d\boldsymbol{y} \leq \beta_1 \\ \qquad \boldsymbol{m}_2^c\hat{\boldsymbol{x}} + \boldsymbol{m}_2^d\boldsymbol{y} \leq \beta_2 \\ \qquad \boldsymbol{y} \geq \boldsymbol{0} \end{array}\right\} \tag{7}$$

and such an optimal solution is called a rational response. Let $R(\boldsymbol{x})$ denote the set of rational responses.

Then, the Stackelberg solution is an optimal solution to the problem

$$\left.\begin{array}{l} \text{minimize } Var[\tilde{\boldsymbol{c}}_1\boldsymbol{x} + \tilde{\boldsymbol{d}}_1\boldsymbol{y}] = \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}^T V_1 \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} \\ \text{subject to } \boldsymbol{y} \in R(\boldsymbol{x}) \\ \qquad A_1'\boldsymbol{x} + A_2'\boldsymbol{y} \leq \boldsymbol{b}' \\ \qquad \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0}, \end{array}\right\} \tag{8}$$

where, for simplicity, we use coefficient matrices and a vector, A_1', A_2' and $\boldsymbol{b}'$, for the coefficients represented by A_1, A_2, $\boldsymbol{m}_1^c$, $\boldsymbol{m}_1^d$, $\boldsymbol{m}_2^c$, $\boldsymbol{m}_2^d$, $\boldsymbol{K}_{1-\alpha}$, β_1 and β_2 of the constraints.

In computation of Stackelberg solutions, it is often assumed that the set of rational responses $R(\boldsymbol{x})$ is a singleton. In the V-model, however, because problem (7) is a convex programming problem and the objective function is also a strictly convex function due to positive definiteness of V_2, the set $R(\boldsymbol{x})$ is a singleton without the assumption.

By replacing the condition of rational responses $\boldsymbol{y} \in R(\boldsymbol{x})$ by the Kuhn-Tucker conditions, we have the following quadratic programming problem including the linear complementarity constraints:

$$\left.\begin{array}{l} \text{minimize } Var[\tilde{\boldsymbol{c}}_1\boldsymbol{x} + \tilde{\boldsymbol{d}}_1\boldsymbol{y}] = \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}^T V_1 \begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} \\ \text{subject to } 2\sum_{j=1}^{n_1} v_{2(n_1+i)j}x_j + 2\sum_{j=n_1+1}^{n_1+n_2} v_{2(n_1+i)j}y_{j-n_1} \\ \qquad +\boldsymbol{\lambda}A'_{2,\cdot i} - \omega_i = 0,\ i = 1, \ldots, n_2 \\ \qquad A_1'\boldsymbol{x} + A_2'\boldsymbol{y} - \boldsymbol{b}' \leq \boldsymbol{0} \\ \qquad \boldsymbol{\lambda}(A_1'\boldsymbol{x} + A_2'\boldsymbol{y} - \boldsymbol{b}') = 0,\ \boldsymbol{\omega}\boldsymbol{y} = 0 \\ \qquad \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0},\ \boldsymbol{\lambda} \geq \boldsymbol{0},\ \boldsymbol{\omega} \geq \boldsymbol{0} \end{array}\right\} \tag{9}$$

Although it is difficult to directly solve problem (9) because problem (9) is not a convex programming problem, it is noted that problem (9) is reduced into a quadratic programming problem with only linear constraints by removing the linear complementarity constraints.

In this paper, we employ the branch-and-bound scheme with branching procedure for the linear complementarity constraints and by iteratively solving quadratic programming problems, we find an optimal solution to problem (9). This method is similar to the Bard and Moore method [2] for solving two-level linear programming problems, in which linear programming problems are iteratively solved.

3 A numerical example

We consider the following two-level linear programming problem in which each of the leader and the follower has only one decision variable x and y, and five constraints:

$$\left.\begin{array}{l} \underset{x}{\text{minimize}}\ z_1(x,y) = \tilde{c}_1 x + \tilde{d}_1 y \\ \text{where } y \text{ solves} \\ \underset{y}{\text{minimize}}\ z_2(x,y) = \tilde{c}_2 x + \tilde{d}_2 y \\ \text{subject to } -x + 3y \leq \tilde{b}_1,\ 10x - y \leq \tilde{b}_2 \\ \qquad\qquad 3x + y \geq \tilde{b}_3,\ x + 2y \geq \tilde{b}_4 \\ \qquad\qquad 3x + 2y \geq \tilde{b}_5,\ x \geq 0, y \geq 0. \end{array}\right\} \tag{10}$$

Means of the random variable coefficients $\tilde{c}_1$, $\tilde{d}_1$, $\tilde{c}_2$ and $\tilde{d}_2$ of the objective functions in the numerical example are −2.0, −3.0, 2.0 and 1.0, respectively, and the covariance matrices with respect to $(\tilde{c}_1, \tilde{d}_1)$ and $(\tilde{c}_2, \tilde{d}_2)$ are $V_1 = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, $V_2 = \begin{bmatrix} 1 & -1 \\ -1 & 6 \end{bmatrix}$. For the right-hand-side constants of the constraints, means, variances and probability levels of satisfaction are shown in Table 1.

Incorporating the chance constrained conditions and employing the proposed formulation, we have the following deterministic two-level quadratic programming problem corresponding to the original problem with random variable coefficients.

$$\left.\begin{array}{l} \underset{x}{\text{minimize}}\ \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ \text{where } y \text{ solves} \\ \underset{y}{\text{minimize}}\ \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ \text{subject to } -x + 3y \leq 47,\ 10x - y \leq 110 \\ \qquad\qquad -3x - y \leq -19,\ -x - 2y \leq -15 \\ \qquad\qquad -3x - 2y \leq -29,\ -2x - 3y \leq -31 \\ \qquad\qquad 2x + y \leq 33,\ x \geq 0,\ y \geq 0 \end{array}\right\} \tag{11}$$

where the last two constraints $-2x - 3y \leq -31$, $2x + y \leq 33$ are those of the expected values of the objective functions of the leader and the follower, respectively. Problem (11) can be transformed to the single level quadratic

Table 1. Random variable coefficients of the constraints

coefficient	$\tilde{b}_1$	$\tilde{b}_2$	$\tilde{b}_3$	$\tilde{b}_4$	$\tilde{b}_5$
mean	50.11	113.15	15.16	13.16	25.63
variance	9.0	36.0	9.0	4.0	16.0
probability	0.85	0.70	0.90	0.70	0.80

Table 2. The Stackelberg solution of each model

model	solution	E-value of the leader	E-value of the follower	variance of the leader	variance of the follower
V-model	(5.1667, 6.8889)	−31	17	266.94	240.25
V-model w/o*	(7, 4)	−26	18	202	89
E-model	(1, 16)	−50	18	802	1505

* V-model w/o means the V-model without the expected value constraints.

programming problem including the linear complementarity constraints as described in the previous section.

As shown in Table 2, although the expected values of the leader and the follower in the E-model are smaller than those of the V-models, the variances of the E-model are considerably larger than those of the V-models. Especially, while the expected values of the follower in the V-models with/without the expected value constraints are almost the same with that of the E-model, the variances are diminished in the V-models to a large extent. It is found that the expected values in the V-model with the expected value constraints can control the level of the expected values, keeping suppressing the variance.

References

1. Bard, J.F., An efficient point algorithm for a linear two-stage optimization problem. Operations Research 38, 556–560, 1983.
2. Bard, J.F. and Moore, J.T., A branch and bound algorithm for the bilevel programming problem. SIAM Journal on Scientific and Statistical Computing 11, 281–292, 1990.
3. Bialas, W.F. and Karwan, M.H., Two-level linear programming. Management Science 30, 1004–1020, 1984.
4. Charnes, A. and Cooper, W.W., Chance constrained programming. Management Science 6, 73–79, 1959.
5. Charnes, A. and Cooper, W.W., Deterministic equivalents for optimizing and satisficing under Chance Constraints, Operations Research 11, 18–39, 1963.
6. Hansen, P., Jaumard, B. and Savard, G., New branch-and-bound rules for liner bilevel programming. SIAM Journal of Scientific and Statistical Computing 13, 1194–1217, 1992.
7. White, D.J. and Anandalingam, G., A penalty function approach for solving bi-level linear programs. Journal of Global Optimization 3, 397–419, 1993.

On Inherited Properties for Vector-Valued Multifunctions

Shogo Nishizawa[1], Tamaki Tanaka[1], and Pando Gr. Georgiev[2]

[1] Graduate School of Science and Technology, Niigata University, Niigata 950–2181, Japan
E-mail: tamaki@m.sc.niigata-u.ac.jp
[2] Department of Mathematics and Informatics, Faculty of Mathematics and Informatics, Sofia University, Sofia, Bulgaria
Current address: Laboratory for Advenced Brain Signal Processing, Brain Science Institute, the Institute of Physical and Chemical Research (RIKEN), Wako, Saitama, Japan

Abstract. In this paper, we consider several kinds of convexity and semicontinuity of multifunctions with respect to ordering cones, and we investigate inherited properties of convexity and semicontinuity through scalarization of multifunctions. We introduce four kinds of scalarizing functions to characterize images of a multifunction by using Tchebyshev scalarization; these (real-valued) scalarizing functions have the same sorts of convexity and semicontinuity which correspond to those of parent multifunctions.

1 Introduction

This paper consists of two parts which are concerned with inherited properties of convexity and semicontinuity. Convexity and lower semicontinuity of real-valued functions are useful properties for analysis of optimization problems, and they are dual concepts to concavity and upper semicontinuity, respectively. These properties are related to the total ordering of R. We consider certain generalizations and modifications of convexity and semicontinuity for multifunctions in a topological vector space with respect to a cone preorder in the target space, which have motivated by [3,4] and studied in [1] for generalizing the classical Fan's inequality. These properties are inherited by the following scalarizing functions;

$$\inf\{h_C(x,y;k) \mid y \in F(x)\} \tag{1}$$

and

$$\sup\{h_C(x,y;k) \mid y \in F(x)\} \tag{2}$$

where $h_C(x,y;k) = \inf\{t \mid y \in tk - C(x)\}$, $F : E \to 2^Y$ is a set-valued map, $C(x)$ a closed convex cone with nonempty interior, x and y are vectors in two topological vector spaces E and Y, respectively, and $k \in \text{int}C(x)$. Note that $h_C(x,\cdot;k)$ is positively homogeneous and subadditive for every fixed $x \in E$ and $k \in \mathrm{int}C(x)$, and that $h_C(x,y;k) \leq 0$ for $y \in -C(x)$. Moreover, we have $-h_C(x,-y;k) = \sup\{t \mid y \in tk + C(x)\}$. This function $h_C(x,y;k)$ has been treated in some papers. Essentially, $h_C(x,y;k)$ is equivalent to the strictly monotonic function defined by Luc [5]. For each $y \in Y$, $h_C(x,y;k) \cdot k$ corresponds the mimimum vector of upper bounds with respect to the cone $C(x)$ restricted to the direction k. Similarly, $-h_C(x,-y;k) \cdot k$ corresponds the maximum vector of lower bounds with respect to the cone $C(x)$ restricted to the direction k. By using such inherited properties as properties of convexity and semicontinuity are inherited from multifunctions into scalarizing functions (1) and (2), four variants of Fan's type inequality for multifunctions have been presented in [2]. Thus, we know that inherited properties are very useful and important.

2 Inherited Properties of Convexity

The aim of this section is to investigate how properties of cone-convexity and cone-concavity are inherited from multifunctions into scalarizing functions (1) and (2). Let E and Y be topological vector spaces and F and $C : E \to 2^Y$ two multivalued mappings. Denote $B(x) = (\mathrm{int}C(x)) \cap (2S \setminus \overline{S})$ (which plays a role of base for $\mathrm{int}C(x)$ without uniqueness), where S is a neighborhood of 0 in Y. We observe the following four types of scalarizing functions:

$$\psi_C^F(x;k) := \sup_{y \in F(x)} h_C(x,y;k), \qquad \varphi_C^F(x;k) := \inf_{y \in F(x)} h_C(x,y;k);$$

$$-\varphi_C^{-F}(x;k) = \sup_{y \in F(x)} -h_C(x,-y;k), \qquad -\psi_C^{-F}(x;k) = \inf_{y \in F(x)} -h_C(x,-y;k).$$

The first and fourth functions have symmetric properties and then results for the fourth function $-\psi_C^{-F}(x;k)$ can be easily proved by those for the first function $\psi_C^F(x;k)$. Similarly, the results for the third function $-\varphi_C^{-F}(x;k)$ can be deduced by those for the second function $\varphi_C^F(x;k)$. By using these four functions we measure each image of multifunction F with respect to its 4-tuple of scalars, which can be regarded as standpoints for the evaluation of the image. To avoid confusion for properties of convexity, we consider the constant case of $C(x) = C$ (a convex cone) and $B(x) = B$ firstly, and $h_C(x,y;k) = h_C(y;k) := \inf\{t \mid y \in tk - C\}$.

To begin with, we recall some kinds of convexity for multifunctions.

Definition 1. A multifunction $F : E \to 2^Y$ is called *C-quasiconvex*, if the set $\{x \in E \mid F(x) \cap (a - C) \neq \emptyset\}$ is convex (or empty) for every $a \in Y$. If $-F$ is C-quasiconvex, then F is said to be *C-quasiconcave*, which is equivalent to $(-C)$-quasiconvex mapping.

Remark 1. The above definition is exactly that of *Ferro type* (-1)*-quasiconvex* mapping in [4, Definition 3.5].

Definition 2. A multifunction $F : E \to 2^Y$ is called (in the sense of [4, Definition 3.6])

(a) *type*-(iii) *C-properly quasiconvex* if for every two points $x_1, x_2 \in E$ and every $\lambda \in [0,1]$ we have either $F(x_1) \subset F(\lambda x_1 + (1-\lambda)x_2) + C$ or $F(x_2) \subset F(\lambda x_1 + (1-\lambda)x_2) + C$.

(b) *type*-(v) *C-properly quasiconvex* if for every two points $x_1, x_2 \in E$ and every $\lambda \in [0,1]$ we have either $F(\lambda x_1 + (1-\lambda)x_2) \subset F(x_1) - C$ or $F(\lambda x_1 + (1-\lambda)x_2) \subset F(x_2) - C$;

If $-F$ is type-(iii) [resp., type-(v)] C-properly quasiconvex then F is said to be *type*-(iii) [resp., *type*-(v)] *C-properly quasiconcave*, which is equivalent to type-(iii) [resp., type-(v)] $(-C)$-properly quasiconvex mapping.

However, there is no relationship between those for types (iii) and (v) in general.

Definition 3. A multifunction $F : E \to 2^Y$ is called (in the sense of [4, Definition 3.7]) *type*-(v) *C-naturally quasiconvex*, if for every two points $x_1, x_2 \in E$ and every $\lambda \in (0,1)$, there exists $\mu \in [0,1]$ such that

$$F(\lambda x_1 + (1-\lambda)x_2) \subset \mu F(x_1) + (1-\mu)F(x_2) - C.$$

If $-F$ is type-(v) C-naturally quasiconvex, then F is said to be *type*-(v) *C-naturally quasiconcave*, which is equivalent to type-(v) $(-C)$-naturally quasiconvex mapping.

Theorem 1.

1. *If $F : E \to 2^Y$ is type-*(v) *C-properly quasiconvex, then*

$$\inf_{k \in B} \psi_C^F(x;k) = \inf_{k \in B} \sup_{y \in F(x)} h_C(y;k)$$

is quasiconvex, and especially $\psi_C^F(x;k)$ is also quasiconvex with respect to variable x for every $k \in \operatorname{int} C$;

2. *If $F : E \to 2^Y$ is type-*(iii) *C-properly quasiconcave, then $\psi_C^F(x;k)$ is quasiconcave with respect to variable x for every $k \in \operatorname{int} C$;*

3. *If $F : E \to 2^Y$ is type-*(v) *C-properly quasiconcave, then $\varphi_C^F(x;k)$ is quasiconcave with respect to variable x for every $k \in \operatorname{int} C$;*

4. *If $F : E \to 2^Y$ is type-*(iii) *C-properly quasiconvex, then $\varphi_C^F(x;k)$ is quasiconvex with respect to variable x for every $k \in \operatorname{int} C$.*

Theorem 2. *If $F : E \to 2^Y$ is C-quasiconvex, then $\varphi_C^F(x;k)$ is quasiconvex with respect to variable x for every $k \in \operatorname{int} C$.*

Theorem 3. *If $F : E \to 2^Y$ is type-*(v) *C-naturally quasiconvex, then $\psi_C^F(x;k)$ is quasiconvex with respect to variable x for every $k \in \operatorname{int} C$.*

Remark 2. When we replace F by $-F$ in the theorems above, it leads to the quasiconvexity (or quasiconcavity) of scalarizing functions $-\psi_C^{-F}$ and $-\varphi_C^{-F}$.

3 Inherited Properties of Semicontinuity

The aim of this section is to investigate how properties of several kinds of cone-semicontinuity is inherited from multifunctions into scalarizing functions. We introduce two types of cone-semicontinuity of multifunctions, which are regarded as extensions of the ordinary lower semicontinuity for real-valued functions; see [3].

Definition 4. Let $\hat{x} \in E$. A multifunction F is called *$C(\hat{x})$-upper semicontinuous* at x_0, if for every $y \in C(\hat{x}) \cup (-C(\hat{x}))$ satisfying with $F(x_0) \subset y + \text{int}C(\hat{x})$, there exists an open $U \ni x_0$ such that $F(x) \subset y + \text{int}C(\hat{x})$ for every $x \in U$.

Definition 5. Let $\hat{x} \in E$. A multifunction F is called *$C(\hat{x})$-lower semicontinuous* at x_0, if for every open V such that $F(x_0) \cap V \neq \emptyset$, there exists an open $U \ni x_0$ such that $F(x) \cap (V + \text{int}C(\hat{x})) \neq \emptyset$ for every $x \in U$.

Remark 3. In the two definitions above, the notions for single-valued functions are equivalent to the ordinary notion of lower semicontinuity of real-valued ones, whenever $Y = R$ and $C(x) = [0, \infty)$. Usual upper semicontinuous multifunction is also (cone-) upper semicontinuous. When the cone $C(\hat{x})$ consists only of the zero of the space, the notion in Definition 5 coincides with that of lower semicontinuous multifunction. Moreover, it is equivalent to the cone-lower semicontinuity defined in [3], based on the fact that $V + \text{int}C(\hat{x}) = V + C(\hat{x})$; see [6, Theorem 2.2].

Proposition 1. ([1, Proposition 3.1]) If for some $x_0 \in E$, $A \subset \text{int}C(x_0)$ is a compact subset and multivalued mapping $W(\cdot) := Y \setminus \{\text{int}C(\cdot)\}$ has a closed graph, then there exists an open set $U \ni x_0$ such that $A \subset C(x)$ for every $x \in U$. In particular C is lower semicontinuous.

We shall say that (F, X), where X is a subset of E, has property (P), if for every $x \in X$ there exists an open $U \ni x$ such that the set $F(U \cap X)$ is precompact in Y, that is, $\overline{F(U \cap X)}$ is compact.

Theorem 4. See [1, Lemma 3.1]. *Suppose that $W : E \to 2^Y$ defined as $W(x) = Y \setminus \text{int}C(x)$ has a closed graph. If F is $(-C(x))$-upper semicontinuous at x for each $x \in E$ and (F, X) satisfies property (P), then $\psi^F|_X$, which is the restriction of*

$$\psi^F(x) := \inf_{k \in B(x)} \sup_{y \in F(x)} h_C(x, y; k)$$

to the set X, is upper semicontinuous. If the mapping C is constant-valued, then ψ^F is upper semicontinuous.

However, we can replace Proposition 1 by another relaxed form as a consequence of itself.

Proposition 2. Assume that there exists a compact subset $D \subset Y$ satisfying (i) $A \subset \text{cone}D$ where $\text{cone}D := \{\lambda x | \lambda \geq 0, x \in D\}$ and (ii) $D \subset \text{int}C(x_0)$ for some $x_0 \in E$. If $W(\cdot) := Y \setminus \{\text{int}C(\cdot)\}$ has a closed graph, then there exists an open set $U \ni x_0$ such that $A \subset C(x)$ for every $x \in U$. In particular C is lower semicontinuous.

Therefore, we consider the following relaxed one instead of property (P): We shall say that (F, X), where X is a subset of E, has property (Q), if for every $x \in X$ there exists a compact set $D(x)$ $(D : X \to 2^Y)$ such that $F(x) \subset coneD(x)$.

Theorem 5. *Suppose that $W : E \to 2^Y$ defined as $W(x) = Y \setminus \text{int}C(x)$ has a closed graph. If F is $(-C(x))$-upper semicontinuous at x for eachx $\in E$ and (F, X) satisfies property (Q), then $\psi^F|_X$, which is the restriction of $\psi^F(x)$ to the set X, is upper semicontinuous. If the mapping C is constant-valued, then ψ^F is upper semicontinuous.*

Theorem 6. See [1, Lemma 3.3]. *Suppose that $W : E \to 2^Y$ defined as $W(x) = Y \setminus \text{int}C(x)$ has a closed graph. If F is $(-C(x))$-lower semicontinuous at x for each $x \in E$ and (F, X) satisfies property (Q), then $\varphi^F|_X$, which is the restriction of*

$$\varphi^F(x) := \inf_{k \in B(x)} \inf_{y \in F(x)} h_C(x, y; k)$$

to the set X, is upper semicontinuous. If the mapping C is constant-valued, then φ^F is upper semicontinuous.

Remark 4. When we replace F by $-F$ in the two theorems above, it leads to the lower semicontinuity of scalarizing functions $-\psi^{-F}$ and $-\varphi^{-F}$.

4 Conclusions

We have established that

(i) if a multifunction has a certain cone-convexity [resp., cone-concavity], that is, C-quasiconvexity, C-properly quasiconvexity, C-naturally quasiconvexity [resp., C- $\cdots$ quasiconcavity], then each scalarizing function is quasiconvex [resp., quasiconcave]. That is one of inherited properties from its parent multifunction;

(ii) if a multifunction has a certain cone-upper semicontinuity [resp., cone-lower semicontinuity], then each scalarizing function is upper semicontinuous [resp., lower semicontinuous]. This is another inherited property from its parent multifunction.

References

1. P. Gr. Georgiev and T. Tanaka (2000). *Vector-valued set-valued variants of Ky Fan's inequality*, Journal of Nonlinear and Convex Analysis, **1(3)**, pp.245–254.
2. P. Gr. Georgiev and T. Tanaka (2001). *Fan's inequality for set-valued maps*, Nonlinear Analysis Theory, Methods and Applications, **47(1)**, pp.607–618.
3. Y. Kimura, K. Tanaka, and T. Tanaka (1999). *On semicontinuity of set-valued maps and marginal functions*, pp.181–188 in Nonlinear Analysis and Convex Analysis —Proceedings of the International Conference (W. Takahashi and T. Tanaka, eds.), World Scientific, Singapore.
4. D. Kuroiwa, T. Tanaka, and T.X.D. Ha (1997). *On cone convexity of set-valued maps*, Nonlinear Analysis, Theory, Methods and Applications, **30(3)**, pp.1487–1496.
5. D. T. Luc (1989). *Theory of Vector Optimization*, Lecture Note in Economics and Mathematical Systems, **319**, Springer, Berlin.
6. T. Tanaka and D. Kuroiwa (1994). *Another observation on conditions assuring* $\text{int}A + B = \text{int}(A + B)$, Applied Mathematics Letters, **7 (1)**, pp.19–22.

Multicriteria Expansion of a Competence Set Using Genetic Algorithm

Serafim Opricovic and Gwo-Hshiung Tzeng

Faculty of Civil Engineering, University of Belgrade, Yugoslavia; in 20001/2002 Opricovic is with the National Chiao Tung University, Hsinchu, Taiwan; e-mail: seropric@yahoo.com; [2]Institute of Management of Technology, National Chiao Tung University, Hsinchu, Taiwan; e-mail: ghtzeng@cc.nctu.edu.tw

Abstract. *Optimal expansion of a competence set, consisting of knowledge, information and skills for a certain decision making problem and its effective solution, is considered in this paper. The competence set expansion is optimized with respect to minimizing costs and time, as well as maximizing efficiency and benefits of expansion. This problem is treated as a multicriteria combinatorial optimization problem. A multicriteria genetic algorithm is developed to solve this optimization problem. The multicriteria measure of closeness to the "ideal" solution is introduced for the fitness assignment. An illustrative example of academic competence set expansion is presented. The results show the applicability of multicriteria genetic algorithm to solve a competence set expansion formulated as a multicriteria optimal route problem.*
Keywords: Competence Set Expansion, Multicriteria, Compromise, Genetic Algorithm

1. Introduction

Ideas, knowledge, information, skills, and every kind of message stored in the human brain, form a "knowledge base" for human thoughts, judgments, actions and reactions. This comprehensive knowledge base is called the "Habitual Domain", a concept proposed by Po-Lung Yu (1990). For each decision problem, there is a *competence set* consisting of knowledge, information and skills for its effective solution (Yu and Zhang 1990). A habitual domain is composed of competence sets. If the existing competence sets of a decision maker are incapable to solve the problem, then the decision maker should expand the competence set related to the problem. Optimal expansion of competence set is considered in this paper as a multicriteria combinatorial optimization problem (optimal route problem).

The Genetic Algorithm(GA) may be applied to network design problems, solving combinatorial optimization problems. A computer code for single-criterion genetic algorithm is presented in Goldberg (1989), and the FORTRAN program for single-criterion GA was developed by Carroll (1999). The first formulations of GAs were essentially single-objective methods, but in the last several years interest in how the multiple objectives can be handled by GA has rapidly increased. Fonseca and Fleming (1998), after reviewing evolutionary approaches, propose that fitness assignment could be interpreted as a multicriteria decision process, and consider the ranking of an arbitrary number of candidates.

This paper focuses on competence set expansion formulated as multicriteria routing problem. Multicriteria expansion of a competence set is discussed in Section 2. In Section 3, multicriteria genetic algorithm is developed to determine a compromise solution that is Pareto optimal. Particular multicriteria fitness is introduced, and a multicriteria problem is transformed into a single-criterion problem solvable using GA. The evolutionary computation, and genetic algorithm as a special case, is chosen as a mathematical tool. A simple numerical experiment in Section 4 illustrates shortly the elements of competence set expansion.

2. Multicriteria Expansion of a Competence Set

There are four basic forms of a competence set, defined as follows:
- the *true competence set* (*Tr(D)*): consisting of ideas, knowledge, skills, attitudes, information and resources that are truly needed to successfully solve problem *D*;
- the *perceived competence set* (*Tp(D)*): the true competence set as perceived by the decision maker (DM);
- the DM's *acquired competence set* (*Ac(D)*): consisting of ideas, knowledge, skills, attitudes, and information that have actually been acquired by the DM;
- the *perceived acquired competence set* (*Ap(D)*): the acquired competence set as perceived by the DM.

The above four forms are all special subsets of the habitual domain (*HD*) for a decision problem *D* (Yu 1990). The gaps between the true competence set (*Tr(D)* or *Ac(D)*) and perceived competence set (*Tp(D)* or *Ap(D)*) are due to ignorance, uncertainty, illusion and wishful thinking. Wisdom and certainty would lead to high quality decision, and illusion and ignorance could lead to low quality decision.

A goal of multicriteria expansion is to reach a truly needed competence set *Tr(D)*. If $Tr(D) \neq Tp(D)$, then the goal of expanding the competence set could be one of them and it should be accepted (maybe DM insists on goal of *Tp(D)*). A simplified situation where $Tr(D)=Tp(D)$ and $Ac(D)=Ap(D)$ (objective perceiving) is considered in this paper.

Optimal expansion of a competence set is an MCDM optimization process with respect to minimizing costs and time, and maximizing efficiency and benefits of expansion. The expansion of a competence set could be represented by a network. The starting network node represents existing competence set and the terminal node is the required competence set *Tr(D)*, the intermediate nodes represent skills and resources needed to reach the terminal node. In the competence set expansion problem, cost is associated with the link *(i,j),* but some benefit (or cost) may be related with node *j*, meaning that benefit is achieved if the competence *j* is realized.

A new formulation of a competence set expansion problem (*CSEP*) is introduced in this paper, as follows.

$$\underset{Nst\in\{Nst\}}{EXT}\{\sum_{i\in Nst} f_{k,ix_i}, k=1,...,K\} \qquad (1)$$

where x_i represents the node within set of nodes that follows node i on the route *Nst* (feasible) from starting to the terminal node, $x_i \in [1, N_i]$; N_i is the number of nodes following node i; the summation is over $i \in Nst$ continuing from starting node to the terminal node (directed walk), and i is the node on the route *Nst*, with an input ordering number within the network; a feasible route *Nst* consists of a set of nodes, $Nst \subset \{x_i, i=1,...,N\}$; $\{Nst\}$ denotes the set of all alternative routes *Nst*; f_{k,ix_i} is the value of *k-th* criterion function for the link (i, x_i), or it is associated with node x_i; *K* is the total number of criterion functions; and *EXT* means that the *k-th* criterion function has to be maximized if it represents benefit, or to be minimized if it represents cost. Each node has its own order number within network, $i=1$ for the starting node and for the terminal node $i=N$. The network is represented by a matrix ***H*** with elements $H_{ij}, j=1,...,N_i, i=1,...,N$, where H_{ij} is the order number of a node that follows node i. The formulation (1) assumes that each node in the network has at least one following node on the route to the terminate node, and there are no "loops" within the network (acyclic). The task is to determine the best route, from the sets $\{x_i, i\in Nst, Nst\in\{Nst\}\}$, in a multicriteria decision making sense.

3. Multicriteria Genetic Algorithm

Genetic algorithms are essentially unconstrained search techniques which require the assignment of a scalar measure of quality (fitness) to the candidate solutions. For a class of constrained problem, a subroutine should be developed for testing feasibility. If the constraints are violated, the current solution is nonfeasible, and thus has no fitness. The fitness assignment is interpreted as a multicriteria decision process.

The *multicriteria genetic algorithm* developed in this paper has the following general steps:

0. Initialization: defining the representation (binary string *s*); population size *M*; and the ending criterion. Defining the fitness function, $fit = F(d(s))$, where *F* is the MCDM aggregated function, including all criteria and decision makers' preferences; and *d* means decoding. Selecting the initial population

$P_1 = \{s_1^1, ..., s_M^1\}$ .Determine $F^* = \max\{F(d(s_m^1)), m = 1, ..., M\}$, where s_m^p denotes m-th individual (a solution) in the p-th population.

1. Determining fitness for all individuals in a current population P_p.

   The procedure for determining multicriteria fitness is presented below.

2. Generating $P_{p+1} = \{s_1^{p+1}, ..., s_M^{p+1}\} \subseteq X_c$, by random performance of genetic operators on random selected individuals from P_p with higher fitness. The execution of GA operators is associated with probability. The probability used in selection is proportional to the fitness, so that a better individual has a better chance to be selected for the evolutionary process.

3. Compute $F_{\max} = \max\{F(d(s_m^{p+1})), m = 1, ..., M\}$.

   If $F_{\max} > F^*$, then $F^* = F_{\max}$ and $x^* = \arg F_{\max}$.

4. If the ending criterion is satisfied, stop the procedure, and x^* represents the optimal solution (or "near-optimum" in some cases); otherwise repeat the steps 1 to 4 with p=p+1 (next generation).

The multicriteria aggregating function $F(d(s))$ should provide sufficient information to guide evolution in GA. The development of function $F(d(s))$ is based on the multicritera compromise ranking method (called VIKOR, Opricovic 1998). The multicriteria measure Q (distance to the "ideal point", representing aggregating function) is introduced to determine multicriteria fitness, $fit_m = 1 - Q_m$, for m-th individual. An individual closer to the "ideal point" has better fitness, and the closest individual represents a "compromise" solution (Yu 1973). The measure Q is determined by the procedure as follows.

$$Q_m = v(S_m - S^*)/(S^- - S^*) + (1-v)(R_m - R^*)/(R^- - R^*), m=1,...,M$$

where: $$S_m = \sum_{k=1}^{K} w_k (f_k^* - f_{km})/(f_k^* - f_k^-),$$

$$R_m = \max_k [w_k (f_k^* - f_{km})/(f_k^* - f_k^-)];$$

w_k are the weights of criteria (given as input data); weights are introduced to express the relative importance of the criteria, and they have no clear economic meaning; f_{km} is the value of the k-th criterion function, determined within a evaluation subroutine, for the solution s_m^p generated by GA; f_k^* and f_k^- are the best and the worst value the k-th criterion function, respectively; S^* and R^* are minimum value of S_m and R_m, respectively; S^- and R^- are maximum values; v is introduced as weight of the strategy of "the majority of criteria" (or "the maximum group utility"), here $v = 0.5$.

The maximal $f_k^{\max}$, and the minimal $f_k^{\min}$ values of all criterion functions, $k=1,\dots,K$ are the input data. These values can be determined by single-criterion optimization, or to be assessed by the analyst. The values of criterion functions for any feasible solution should be within given intervals. Also, S^-, S^*, R^-, R^* are needed as input data. These data could be given by a decision maker (firm data), or to be given as initial assessment. With "firm data" the results are obtained within one GA run. With "assessed data" entered for the first GA run, the new values for $f_k^{\max}$, $f_k^{\min}$, $k=1,\dots,K$, and S^-, S^*, R^-, R^* are computed for the entire population size, and used as input data for the second GA run. An extreme (possible) value for $f_k^{\max}$, $f_k^{\min}$, $k=1,\dots,K$, and $S^-=R^-=1$, $S^*=R^*=0$, may be given as "assessed data" for the first GA run.

The compromise solution could be accepted by the decision makers because it provides a maximum "group utility" for the "majority" (with measure *S* representing "concordance"), and a minimum of individual regret by the "opponents" (with measure *R* representing "discordance").

4. Illustrative Example

As an illustrative example, the competence set expansion for water resources management could be formulated. The illustrative data are taken from the course selection guide for Civil Engineering Department, FAMU/FSU College of Engineering, Tallahassee, Florida. An input is the list of courses for developing knowledge in water resources management. The prerequisite courses for the main courses also have their own prerequisites, as the courses for competence set expansion.

The original network for this example is formulated including all courses, with Calculus I as the starting node. The node with more than one competence is a "compound" node. There are alternative routes within such a network, from the starting node to the terminal node *Tr(D)*. The terminal node represents the needed competence set for water resources management. The loop may be transformed in oriented link by adding a new node (artificial) and a new link. The compound node is noted with the code of first course, and the values of criterion functions for such node are the integral for the compound node.

The criteria could be : f_1 - credit hours, f_2 - design credits, f_3 -time (needed for expansion), and f_4 -expected efforts. The criterion functions f_1 and f_2 are associated with nodes. The criterion function f_3 represents the total number of courses within one alternative route from starting node to the terminal node *Tr*. The criterion function f_4 represents expected efforts to pass the courses on a given route, and the values are subjectively evaluated. The total value of the criterion functions f_1 and f_2 are to be maximized for the optimal route, and f_3 and f_4 are to be minimized.

The multicriteria optimal solution (compromise) is determined by program EMCO (Evolutionary Multicriteria Combinatorial Optimization) based on the methodology presented in this paper. Population size of 8 individuals, and 200 generations were the GA parameters for this example. Multicriteria optimization is performed for two preference "scenarios" represented by the weights values, the CSE_1 (Competence Set Expansion) solution for "preference on hydraulic engineering" and the CSE_2 solution representing the "preference on credit hours". The DM's preference has great impact on the MCDM solution. These solutions illustrate two ways of competence set expansion, achieving the goal of acquiring knowledge for water resources management. The proficiency level of the achieved competence set is not considered in this example.

5. Conclusions

Competence set expansion for solving a certain problem is considered as multicriteria decision making, and it is formulated as a multicriteria optimal route problem. Using the newly developed multicriteria genetic algorithm, a compromise solution is determined. An illustrative example shows that the proposed method and algorithm can solve a competence set expansion problem.

New model of a competence set expansion may be considered as a contribution of this paper.

The application of genetic algorithm is efficient in solving combinatorial optimization problems, although the convergence and global optimum are a task for future research on developing genetic algorithms. Some of the parameters in the model of competence set expansion may not be crisp, and there is a need to develop a fuzzy multicriteria genetic algorithm that includes fuzzy logic.

Acknowledgments-- This paper is partly a result of the project NSC90-2811-H-009-001, which was supported by the National Science Council of Taiwan.

References

Carroll DL (1999) FORTRAN Genetic Algorithm (GA) Driver. WWW page, http://www.staff.uiuc.edu/~carroll/ga.html.

Fonseca C, Fleming P (1998) Multiobjective Optimization and Multiple Constraint Handling with Evolutionary Algorithms - Part I: A Unified Formulation. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 28(1): 26-37.

Goldberg D (1989) Genetic Algorithms in Search, Optimization, and Machine Learning Addison-Wesley Publishing Company, Inc., New York.

Opricovic S (1998) Multicriteria Optimization in Civil Engineering Faculty of Civil Engineering, Belgrade.

Yu PL (1973) A Class of Solutions for Group Decision Problems. Management Science 19(8): 936-946.

Yu PL (1990) Forming Winning Strategies – An Integrated Theory of Habitual Domains. Springer-Verlag, Heidelberg.

Yu PL, Zhang D (1990) A Foundation for Competence Set Analysis. Mathematical Social Sciences 20(3): 251-299.

Comparing DEA and MCDM Method

Serafim Opricovic and Gwo-Hshiung Tzeng

Faculty of Civil Engineering, University of Belgrade, Yugoslavia; in 20001/2002 he is with the National Chiao Tung University, Hsinchu, Taiwan; e-mail: seropric@yahoo.com; [2]Institute of Management of Technology, National Chiao Tung University, Hsinchu, Taiwan; e-mail: ghtzeng@cc.nctu.edu.tw

Abstract. *Data Envelopment Analysis (DEA) introduces a model for weights determination maximizing efficiency of the decision-making units. The primary focus of the DEA model is to compare decision-making units (alternatives) in terms of their efficiency in converting inputs into outputs. A multicriteria decision making (MCDM) method uses a common set of weights that express a decision maker's preferences. In contrast, the DEA does not provide a common set of weights that could express the preferences of a decision maker. A comparison of DEA and MCDM shows that DEA resembles MCDM, but the results differ. In spite of these differences, DEA could be used as a supplement for screening alternatives within MCDM.*
Keywords: Data envelopment analysis, Multicriteria decision making, Comparison

1. Introduction

Data Envelopment Analysis (DEA), developed by Charnes et al.(1978), is a linear programming technique used to estimate the relative efficiency of decision-making units (DMUs), considering the multiple inputs that they consume, and multiple outputs that they produce. A standard formulation of DEA creates a separate linear programming model for each DMU, in which the unknown variables are the weights associated with inputs and outputs. The basic result of DEA is an envelopment surface (efficient frontier) consisting of the "best practice" decision-making units, as well as an efficient measure that reflects the distance from each DMU to the frontier.

A relationship between DEA and multicriteria decision making was considered by Stewart (1996), who concluded that the fields of MCDM and DEA have developed, to a large extent, independently of each other. The criteria in MCDM can be divided into costs (inputs) and benefits (outputs) which give the methodological connection between DEA and MCDM. A DMU within DEA is usually called an alternative within MCDM. Charnes et al. (1978) recognized the difficulty in seeking a common set of weights to determine relative efficiency, and the difficulty in bounding the weights was considered by Stewart (1996). Chiang and Tzeng

(2000) applied a multiple objective programming method in determining a common set of weights. They considered every DMU's efficiency as one objective function to be maximized, and the solution is determined by maximizing minimum DMUs' efficiency.

The compromise ranking method (called VIKOR) has been introduced as a useful technique to implement within MCDM (Opricovic 1998). Assuming that each alternative is evaluated according to each criterion function, the compromise ranking could be performed by comparing the measure of closeness to the ideal alternative. The compromise programming method (Yu 1973) introduced L_p *-metric* as an aggregating function, and the development of VIKOR method started with the metric introduced by Duckstein and Opricovic (1980). The main VIKOR result is the ranking list of alternatives, and the compromise solution with the "advantage rate". Ranking by VIKOR may be performed using different values for criteria weights, and then analyzing the impact of criteria weights on proposed compromise solution. The compromise solutions could be the base for negotiation, involving the decision makers' preferences by criteria weights.

An intention of this our paper is to compare the procedural basis of a DEA method (CCR model) and a MCDM method, VIKOR, as well as to apply DEA as a supplement within multicriteria decision making. DEA provides an efficiency measure that does not rely on the application of a common weighting of the inputs and outputs. In contrast, a multicriteria decision making approach is based on the assumption that a common set of weights must be applied across all units (alternatives). Numerical experiment showed difference between results obtained by DEA and MCDM, and the DEA results were not so useful within MCDM as we expected. Finally, it is concluded that DEA does not provide a common set of weights that could express the preferences of a decision maker. DEA is an approach different from the MCDM method, VIKOR.

2. Comparison of DEA and VIKOR

DEA introduces a linear programming model for weights determination, individually maximizing the efficiency of the decision-making units (DMUs). Therefore the DMUs cannot be ranked with these weights, which vary from unit to unit. Any common set of weights based on the DEA results has no relation with the preference of decision maker who is competent to assess the relative importance of the criteria within MCDM. This is because DEA provides an efficiency measure that does not rely on the application of a common weighting of the inputs and outputs. On the contrary, a multicriteria decision making approach is based on the assumption that a common set of weights must be applied across all alternatives (decision-making units).

We have analyzed similarity of DEA and the VIKOR method of MCDM in order to find a way of using DEA results for multicriteria decision making, particularly in assessing criteria weights. However, we found that the DEA approach was different from a MCDM method and the DEA results were not as useful within MCDM as expected.

The fields of MCDM and DEA have developed, to a large extent, independently of each other (Doyle and Green 1993; Stewart 1996). However, decision-making units, inputs and outputs in DEA can be considered as alternatives, costs and benefits in MCDM, respectively. This relationship provides a basis for a comparison of DEA and the VIKOR method of MCDM. From this standpoint a similarity is evident, although there are essential differences within.

We did several numerical experiments that compare DEA and VIKOR methods, and the findings are summarized below.

- *Efficiency and Pareto optimality*: The concept of Pareto optimality is the core of DEA and MCDM. However, the frontier by DEA is in the space of output/input ratios, and Pareto optimality in MCDM is considered in criteria space. This is why the positions of DMUs (alternatives) are different in these spaces.
- *Decision criterion*: In DEA this is the ratio of multiple outputs and multiple inputs, while in VIKOR it is the aggregating function (distance function) of all criteria. Any DMU, that performs the best on one particular ratio of an output to an input, is found to be efficient by DEA; while a noninferior solution in MCDM is any DMU with at least one input or one output as the best.
- *Solution*: The set of efficient units determined by DEA has no relationship with noninferior solutions within MCDM, whereas the compromise solution by VIKOR is a noninferior solution (Pareto-optimal). An efficient unit is a noninferior solution in the space of output/input ratios considered by DEA. A noninferior solution within MCDM could be inefficient unit by DEA. An efficient unit determined by DEA could be the best compromise solution by VIKOR, although an inefficient unit by DEA also could be the best compromise solution by VIKOR. In many cases, efficient units by DEA are highly raked by VIKOR, and very inefficient units by DEA are given low rankings by VIKOR, although the exception could be the alternative with the extreme value of certain criterion.
- *Weights*: The values of weights (u,v) determined by DEA are not related to the decision makers' preference; whereas in MCDM the criteria weights are assessed or given by decision makers.
- *Usefulness*: DEA determines the efficient DMUs and generates potential improvements for inefficient DMUs. In contrast, the VIKOR method ranks alternatives by comparing the measure of closeness to the ideal alternative, and then selects the best (compromise) alternative from a set of alternatives in the presence of conflicting criteria.

In spite of these differences, DEA could be a preprocess in MCDM, providing a substantial screening of alternatives for MCDM. Because DEA determines the efficient DMUs without any information of the relative importance of inputs and outputs, it could be a useful tool in MCDM, particularly when a decision maker is not able to express a preference at the beginning of system design or planning. However, DEA can not replace MCDM in selecting the best (compromise) solution for the MCDM problem.

The potential improvements (target and Δf = *target - actual*) for inefficient units by DEA (obtainable by *Frontier Analyst* software 2000) show how an inefficient DMU needs to decrease its inputs or increase its outputs (to the target values) in order to become efficient. This is very useful result within DEA application, but it is of less interest within MCDM.

3. Numerical Experiment

Previous studies of hydropower potential for the Drina River, in former Yugoslavia, have selected potential dam sites for reservoirs to provide hydropower. In addition, comprehensive analysis was required to resolve conflicting technical, social and environmental features. The reservoir systems consist from one to four reservoirs. The alternatives were generated by varying two system parameters, dam site and dam height. Six alternatives were selected and evaluated according to the following eight criteria: f_1 - Profit (10^6 Dinar, Yugoslav currency), f_2 - Costs (10^6 Dinar), f_3 - Total energy produced (GWh/year), f_4 - Peak energy produced (GWh/year), f_5 - Number of homes to be relocated, f_6 - Area flooded by reservoirs (ha), f_7 - Number of villages to displace (even partially), f_8 - Environmental protection (grades 1 to 5). The values of criterion functions are obtained by a comprehensive study of this reservoir system on Drina River. The multicriteria optimization task is to maximize the criterion functions f_1, f_3, f_4, and f_8, and to minimize functions f_2, f_5, f_6, and f_7.

DEA application. The decision-making units (DMUs) are : $A_1, A_2, A_3, A_4, A_5, A_6$. The inputs are: f_2 (costs), f_5, f_6, and f_7 (social impacts, land resources). The outputs are: f_1 (profit), f_3 and f_4 (energy produced), and f_8 (lower environmental impact evaluated higher).

Linear programming problem DEA-LP could be solved by a linear programming program package, such as LINDO. Efficient DMUs (alternatives) are: A_3, A_5, A_6 (*Eff* = 1), and inefficient DMUs are: A_1, A_2, A_4 (*Eff* < 1). Ranking based on efficiency is as follows: $A_3 \approx A_5 \approx A_6, A_1, A_2, A_4$.

This application of DEA indicates the set $\{A_3, A_6, A_5\}$ as good alternatives, selecting them as candidates for the best solution within MCDM. This is DEA's main usefulness for multicriteria decision making.

VIKOR application. Alternatives $A_1, A_2, A_3, A_4, A_5, A_6$ are ranked using the VIKOR method with five sets of weights values. The even criteria weights, unnormalized values W1 = $\{ w_i = 1, \forall i \}$, represent indifference of the decision maker. The criteria weights W2 = $\{ w_i = 2, i = 1,2,3,4;\ w_i = 1, i = 5,6,7,8 \}$ express an economic preference. The weights W3 = { $w_i = 1, i =$ 1,2,3,4; $w_i = 2, i = 5,6,7,8$} express preference for social attributes and environment, and W4 = { $w_i = 1, i = 1,2,3,4$; $w_i = 3.2, i = 5,6,7,8$} emphasizes more social criteria.

The ranking results indicate alternative A_5 as the best ranked. It has a good advantage for the weight sets W1 and W2. With the weights W3and W4 the compromise sets are obtained $\{A_5, A_3, A_6\}$, $\{A_3, A_5, A_6\}$, respectively. In these cases the first ranked alternative has no advantage to be a single solution. If the weights of social criteria are increased, such as W4, the alternative A_3 moves to the first place.

A comparison of DEA and VIKOR results. The weight sets W1, W2, W3, and W4 were proposed by the decision maker for the Drina project in order to analyze the preference stability of the compromise solution.

The preference of a decision maker (DM) regarding the relative importance of criteria is not included within a DEA application. This preference could strongly affect the selection of an alternative as a final (preferred) solution. Inclusion of DM preference is one of the main differences between DEA and MCDM.

All six alternatives $A_1,\ldots,A_6$ are noninferior solutions within MCDM. Three of these alternatives, $\{A_3, A_5, A_6\}$, are efficient by DEA, and three are inefficient DMUs, $\{A_1, A_2, A_4\}$. Efficient DMUs $\{A_3, A_5, A_6\}$ are highly ranked by VIKOR, and inefficient DMUs $\{A_1, A_2, A_4\}$ are mainly low ranked. Alternative A_5 (or A_3) is the best ranked by VIKOR.

Alternative A_3 has the best ratios of f_8 / f_5 and f_8 / f_7, it is an efficient DMU, and it is the best compromise solution by VIKOR for certain weights. Alternative A_4 is inefficient by DEA since it has no single best ratio output/input. However, it is the best according to f_1, f_3, and f_4 (all outputs), and it is the best as ranked by VIKOR with the weights W5 = {3, 1, 3, 3, 1, 1, 1, 2}.

Discussion and Proposed Solution. The results by both methods, DEA and VIKOR, indicate the set $\{A_3, A_5, A_6\}$ as good alternatives. As an alternative for a final solution, alternative A_5 could be considered the best compromise. Alternative A_5 is closer to the ideal according to the "economic" criteria f_1, f_2, f_3, f_4. The alternative A_3 has an additional "defect" in that it is more expensive, although it would be preferred from the social point of view. It may be concluded that three alternatives $\{A_3, A_5, A_6\}$ indicated as good solutions. The alternatives A_5 and A_6 are similar three-reservoir systems, where two reservoirs are the same. The alternative A_3 is a system of four small reservoirs. The decision makers for the Drina project prefer alternative A_5, which could be developed in two phases. The first phase develops the system of two reservoirs, and the second phase adds the third reservoir, with a different dam site that could be analyzed later (alternatives A_5 and A_6).

4. Conclusions

The focus of the VIKOR method is that of selecting from a set of alternatives in the presence of conflicting criteria by analyzing the criteria space. The weights in MCDM do not have a clear economic significance, but their use provides the opportunity to model actual aspects of decision making.

The primary focus of DEA model is that of comparing decision-making units (alternatives) from the point of view of their efficiency in converting inputs into outputs. DEA introduces a model for weights determination individually maximizing efficiency of the decision making units. Therefore the DMUs cannot be ranked with these weights that vary from unit to unit. However, DEA could be a preprocess in MCDM, providing screening of alternatives, particularly when the decision maker is not able to express preferences at the beginning of system design or planning.

Finally, DEA resembles MCDM, but it provides different results. In spite of these differences, DEA could be useful in screening alternatives, and identifying efficient units as candidates for the best solution within MCDM. DEA and MCDM method both are helping decision makers to decide the preferred solution. Further research on DEA modifications could bring DEA closer to MCDM.

Acknowledgments. This paper is partly a result of the project NSC90-2811-H-009-001, which was supported by the National Science Council of Taiwan, and the project Hydropower Systems Optimization, which was supported by the Ministry of Science, Serbia.

References

Charnes A, Cooper W, Rhodes E (1978) Measuring the efficiency of decision making units. European Journal of Operational Research **2**: 429-444.

Ciang CI, Tzeng GH (2000) A Multiple Objective Programming Approach to Data Envelopment Analysis. In: Shi Y. and Zeleny M. (eds). New Frontiers of Decision Making for the Information Technology Era, World Scientific, Singapore.

Doyle J, Green R (1993) Data Envelopment Analysis and Multiple Criteria Decision Making. Omega **21**: 713-715.

Duckstein L, Opricovic S (1980) Multiobjective Optimization in River Basin Development. Water Resources Research **16**: 14-20.

Frontier Analyst (2000). Banxia Software Ltd, Glasgow.

Opricovic S (1998) Multictiteria Optimization of Civil Engineering Systems. Faculty of Civil Engineering, Belgrade.

Stewart T (1996) Relationships between Data Envelopment Analysis and Multicriteria Decision Analysis. Journal of the Operational Research Society **47**: 654-665.

Yu PL (1973) A Class of Solutions for Group Decision Problems. Management Science **19**: 936-946.

Linear Coordination Method for Multi-Objective Problems

Busaba Phruksaphanrat [1] and Ario Ohsato [2]

[1] Graduate school of Engineering, Nagaoka University of Technology

[2] Department of Management and Information Systems Science,

Nagaoka University of Technology

At present, the most commonly used satisficing method for multi-objective linear programming (MOLP) is goal programming (GP) based methods but these methods do not always generate efficient solutions. Recently, an efficient GP-based method, which is called reference goal programming (RGP), has been proposed. However, it is limited to only a triangular preference. The more flexible preferences such a convex polyhedral type is preferred in many practical problems. In this research, a satisfactory effective linear coordination method for MOLP problems with convex polyhedral preference functions is proposed. It can be solved by existing linear programming solvers and can find all of the efficient solutions, which satisfy decision maker's requirements. The convex polyhedral function enriches the existing preferences for efficient methods and increases the flexibility in designing preferences.

1 Introduction

Fundamental to a multi-objective linear programming (MOLP) problem is *Pareto optimal concept*, which is also known as an *efficient solution* or a *nondominated solution*. The efficient solution of the MOLP problem is one where any improvement of one objective function can be achieved only at the expense of another [13]. However, the efficient solutions of the real-world problems are noncomparable and so large for the practical uses. So, the need of considering additional information such a preference function arises. Normally, the preference modeling technique is applied to goal programming (GP) based methods [1-4], which are MOLP under satisficing concept [10-12]. The solution from GP-based method may not be the efficient solutions. Recently, reference goal programming (RGP) [6,7,11] is proposed for finding the efficient solution of an MOLP problem. This method expresses the reference point method (RPM) [12] by GP. It always guarantees the efficiency of the solution, which is different from typical GP formulations. However, only a triangular preference is considered. Interval preference structures, increasing in preferences, decreasing in preferences, or other piecewise linear preferences are more flexible [4]. These kinds of preference structures can be collectively called *convex polyhedral* [9] preference functions. GP with these preference functions have been used in many practical problems [4,5]. However, it can find only satisficing solutions, which are not always efficient. This weakness has led to the development of the satisfactory efficient linear coordination method in this research.

2 Lexicographic Models

Consider a minimization problem with K objective functions as follows:

$$\min\{[f_1(\mathbf{x}), f_2(\mathbf{x}),\ldots, f_K(\mathbf{x})] : \mathbf{x} \in Q\} \tag{2.1}$$

where $\mathbf{x}$ denotes a vector of decision variables to be selected with in the feasible set Q; $\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]^T$, and $f_k(\mathbf{x})$ is the kth objective function.

RPM is the technique where the decision maker specifies preferences in terms of reference levels, which can be modeled by the following problem:[6,7]

$$\text{lex min}\left\{\left[\max_{1\le k\le K}\{g_k(l_k, y_k)\}, \sum_{k=1}^{K} g_k(l_k, y_k)\right] : \mathbf{x} \in Q\right\} \tag{2.2}$$

where y_k denotes the mathematical expression of kth objective, ($y_k = f_k(\mathbf{x})$), l_k denote reference levels, and $g_k : R^2 \to R$, for $k = 1,2,\ldots,K$, are the individual achievement functions measuring actual achievement of the kth objective with respect to the corresponding reference level, l_k.

The advantage of the above lexicographic model is that it allows the DM to generate all efficient solutions. Recently, Ogryczak [7] has proposed 4 priority levels of RGP model. The corresponding RGP model always guarantees the efficiency of solutions. However, only a triangular preference can be used. Various preference functions g_k provide a wide modeling environment for measuring individual achievements. The piecewise linear preference functions which is called convex polyhedral preference functions as shown in **Fig.1.** should be employed.

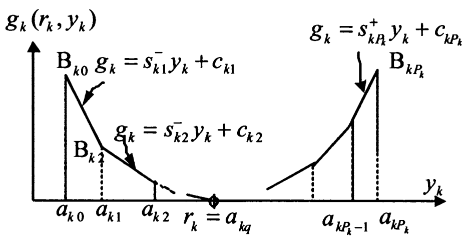


Fig. 1. The convex polyhedral preference function of the kth objective

Convex preference function can be mathematically represented by

$$g_k(r_k, y_k) = \begin{cases} s_{k1}^- y_k + c_{k1}, & \text{if } a_{k0} \le y_k < a_{k1}, \\ s_{k2}^- y_k + c_{k2}, & \text{if } a_{k1} \le y_k < a_{k2}, \\ \vdots & \\ s_{kq}^- y_k + c_{kq}, & \text{if } a_{kq-1} \le y_k < a_{kq} \text{ ; where } a_{kq} = r_k \\ s_{kq+1}^+ y_k + c_{kq+1}, & \text{if } a_{kq} \le y_k < a_{kq+1} \\ \vdots & \\ s_{kP_k}^+ y_k + c_{kP_k}, & \text{if } a_{kP_{k-1}} \le y_k < a_{kP_k}, \end{cases} \tag{2.3}$$

where

r_k be the aspiration level or the reference level, l_k,

a_{kd} be the dth breakpoint of $g_k(r_k, y_k)$, $d = 0, 1, \ldots, P_k$, $k = 1, \ldots, K$,

s_{kd}^- and s_{kd}^+ be the slope of the line segment in the range (a_{kd-1}, a_{kd}) of the negative and the positive side of r_k,

c_{kd} be the y-intercept of the corresponding line segment,

s_{kq}^-, c_{kq} and s_{kq+1}^+, c_{kq+1} are the corresponding slope and the y-intercept of the line segment of the negative and the positive side of a_{kq}, $a_{kq} = r_k$.

3 Efficient Linear Coordination Method Based on Convex Cone Concept

3.1 The concept of convex cone

In the concept of convex cone, it is possible to find any vector $\vec{Z}$ in the convex cone V by the following equation:

$$V = \left\{ \vec{Z} \middle| \vec{Z} = \sum_{k=1}^{K} \lambda_k \vec{D}_k, \lambda_k \ge 0, k = 1,2,\ldots,K \right\} \tag{3.1}$$

where $\vec{D}_k$ are vectors from an extreme point, E to some points on considering space and λ_k are coefficients related to $\vec{D}_k$. With the additional constraint,

$$\sum_{k=1}^{K} \lambda_k \le 1 \tag{3.2}$$

a bounded convex cone V_B can be formed within the convex cone V. This bounded convex cone means a *convex polyhedral* [9].

The convex cone concept in Eq.(3.1) and Eq.(3.2) can be used to form convex polyhedral preference functions in MOLP problems by linear functions so it can be called *a linear coordination method.*

3.2 Formulation of linear coordination method based on convex cone concept

MOLP with convex polyhedral preference functions as shown in **Fig.1.** can be formulated. The detail of formulation is shown in Ref.[9].

Let b_{kd}^- and b_{kd}^+ be deviational constants in the negative and positive sides of the aspiration level, r_k on y_k axis, where $d = 0,1, \ldots, q$ for b_{kd}^-, $d = q+1, \ldots, P_k$, for b_{kd}^+ and g_{kd} be the normalized value of function $g_k(r_k, y_k)$ at breakpoint B_{kd}.

The formulation of a convex polyhedral preference function for a single objective problem based on convex cone concept can be shown as

$$\min \sum_{d=1}^{P_k} g'_{kd}\,\lambda_{kd}\,, \tag{3.3}$$

subject to
$$f_k(\mathbf{x}) + \sum_{d=0}^{q} b_{kd}^- \lambda_{kd} - \sum_{d=q+1}^{P_k} b_{kd}^+ \lambda_{kd} = r_k\,,$$
$$\sum_{d=1}^{P_k} \lambda_{kd} \le 1\,,$$
$$\mathbf{x} \in Q;\ \mathbf{x}, \lambda_{kd} \ge 0,, d = 0,1,\ldots, P_k.$$

Both additive model and minmax model can be applied. However, they are based on the satisficing concept, which cannot ensure the efficiency of solutions.

3.3 Efficient linear coordination method based on convex cone concept

From(3.3), Lexicographic RPM formulation stated in (2.2) can be considered as,

$$(f_k(\mathbf{x}) - r_k) = -\sum_{d=1}^{q} b_{kd}^- \lambda_{kd} + \sum_{d=q+1}^{P_k} b_{kd}^+ \lambda_{kd}\,, \tag{3.4}$$

with additional constraint $\sum_{d=1}^{q} b_{kd}^- \lambda_{kd} \times \sum_{d=q+1}^{P_k} b_{kd}^+ \lambda_{kd} = 0$ for all k. Then, the efficient linear coordination method can be formulated as

$$\operatorname{lex\,min}\left[\max_{1\le k\le K}\Big(\sum_{d=q+1}^{P_k} g'_{kd}\lambda_{kd}\Big), \sum_{k=1}^{K}\Big(\sum_{d=q+1}^{P_k} g'_{kd}\lambda_{kd}\Big), \max_{1\le k\le K}\Big(-\sum_{d=1}^{q} g'_{kd}\lambda_{kd}\Big), \sum_{k=1}^{K}\Big(-\sum_{d=1}^{q} g'_{kd}\lambda_{kd}\Big)\right] \tag{3.5}$$

subject to
$$f_k(\mathbf{x}) + \sum_{d=0}^{q} b_{kd}^- \lambda_{kd} - \sum_{d=q+1}^{P_k} b_{kd}^+ \lambda_{kd} = r_k, k = 1,2,\ldots, K,$$
$$\sum_{d=1}^{P_k} \lambda_{kd} \le 1, k = 1,2,\ldots, K,$$
$$\mathbf{x} \in Q;\ \mathbf{x}, \lambda_{kd} \ge 0,\ k = 1,\ldots, K,\ d = 0,1,\ldots, P_k.$$

Note that $\sum_{d=1}^{q} b_{kd}^- \lambda_{kd} \times \sum_{d=q+1}^{P_k} b_{kd}^+ \lambda_{kd} = 0$ is directly put into the model in Eq.(3.5) by considering the negative and positive vectors from the reference level lexicographically so this constraint can be omitted. The effective linear coordination model is more advantage than existing methods because it can generate the efficient solutions. The effectiveness is illustrated in the example.

4. Numerical Example

We compare the results between linear coordination method and GP-based methods (Jones and Tamiz' s formulation [10]) of the following example.

$$\min \quad f_1(x) = x_1, \tag{4.1}$$

$$\min \quad f_2(x) = x_2,$$

$$\text{subject to} \quad 3x_1 + 4x_2 \geq 30,\ x_1 \geq 2, x_2 \geq 3.$$

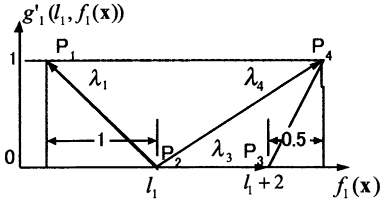


Fig. 2. Preferences function of $f_1(x)$

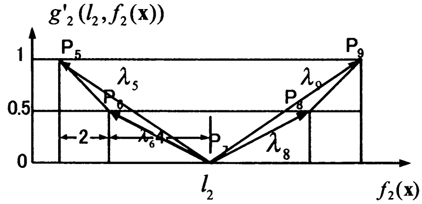


Fig. 3. Preferences function of $f_2(x)$

Let l_1 and l_2 represents the aspiration or reference levels of objective 1 and objective 2 accordingly. Effective linear coordination model can be formulated as:

$$\text{lexmin}\left[\max(\lambda_4, 0.5\lambda_8 + \lambda_9), (\lambda_4 + 0.5\lambda_8 + \lambda_9), \max(-\lambda_1, -\lambda_5 - 0.5\lambda_6), (-\lambda_1 - \lambda_5 - 0.5\lambda_6)\right] \tag{4.2}$$

$$\text{subject to} \quad x_1 + \lambda_1 - 2\lambda_3 - 2.5\lambda_4 = l_1,$$

$$\lambda_1 + \lambda_3 + \lambda_4 \leq 1,$$

$$x_2 + 6\lambda_5 + 4\lambda_6 - 4\lambda_8 - 6\lambda_9 = l_2,$$

$$\lambda_5 + \lambda_6 + \lambda_8 + \lambda_9 \leq 1,$$

$$3x_1 + 4x_2 \geq 30,\ x_1 \geq 2, x_2 \geq 3.$$

Table 4.1 Optimal solutions from different methods solving by linear programming

l_1, l_2	WGP	Minmax GP	Proposed method
1,8	*[2,3],8	*[2,3],8	2,6
2,8	*[2,4],8	*[2,4],8	2.73,5.45
2,5	![3.33,4],5	![3.33,4],5	3.82,4.64
3,4	![4.66,5],4	![4.66,5],4	4.91,3.82
4,4	![4.66,5],4	![4.66,5],4	3.77,4.67
4,5	*[4,6],5	*[4,6],5	3,5.25
5,4	*[5,7],4	*[5,7],4	4,4.5
5,5	*[5,7],5	*[5,7],5	6,3

* Non-efficient solution

! The solutions contain efficient solution and non-efficient solutions

As shown in **Table 4.1**, the proposed method can find all of the efficient solutions, while existing method could not find all of the efficient solutions.

5 Conclusion

The linear coordination method for convex polyhedral preference functions of an MOLP problem is proposed in this research. Convex cone concept and the existing lexicographic model of RPM are applied to formulate the efficient linear coordination method, which can be easily solved by linear programming. The solution from this method is always efficient and also close to the decision maker's requirements. This method has the better solution than the existing methods in the sense of the Pareto optimality. Furthermore, the flexibility in designing preference functions is also enhanced. This method can generate a single efficient solution for the decision maker by only one reference point.

References

1. Charnes A, Cooper WW (1977) Management Models and Industrial Applications of Linear Programming Vol I, John Wiley & Sons Inc., Americas
2. Charnes A, Cooper WW (1977) Goal Programming and Multiple Objective Optimizations Part I. European Journal of Operational Research 1: 39-54
3. Ignizio JP (1976) Goal Programming and Extensions, Lexington Books, Americas
4. Jones DF and Tamiz M (1995) Expanding the Flexibility of Goal Programming via Preference Modeling Techniques. Omega 23: 41-48
5. Kvanli AH (1980) Financial Planning using Goal Programming Omega 8: 207-218
6. Ogryczak W (1994) A goal programming model of the reference point method. Annals of Operation Research 51: 33-44
7. Ogryczak W (2001) On goal programming formulations of the reference point method. Journal of Operational Research Society 52: 691-698
8. Ogryczak W (2001) Comment on properties of the minmax solutions in goal programming. European Journal of Operational Research 132: 17-21
9. Phruksaphanrat B and Ohsato A (2002) Effective linear computational method for nonlinear optimization with a convex polyhedral objective function and linear constraints. Journal of Advanced Computational Intelligence 6: 7-18.
10. Romero C (1991) Handbook of Critical Issues in Goal Programming, Pergamon Press, Great Britain.
11. Tamiz M, Jones DF, Romero C (1998) Goal programming for decision making: An overview of the current state-of-the–art. European Journal of Operational Research 111: 569-581
12. Wierzbicki AP (1982) A Mathematical basis for satisficing decision making. Mathematical Modeling 3: 391-405
13. Zeleny M (1982) Multiple Criteria Decision Making, McGraw-Hill, Americas

Experimental Analysis for Rational Decision Making by Aspiration Level AHP

Kouichi Taji, Junsuke Suzuki, Satoru Takahashi, and Hiroyuki Tamura

Department of Systems and Human Science, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, JAPAN

Abstract. Recently, Triantaphyllou reported that, when evaluating binomial relation by the original analytic hierarchy process (AHP), the ranking of alternatives is different from that obtained when evaluating all alternative simultaneously, even if the preference transitivity holds and if all pairwise comparison matrices are completely consistent. In this paper, we show that the aspiration level AHP, one of the AHP modifications proposed by Tamura et al., does not cause such irrational ranking for completely consistent case. We also estimate how often irrational rankings occur in practical cases by experimental analysis.

1 Introduction

AHP (Analytic Hierarchy Process) proposed by Saaty [1] is one of decision making tools. To apply to a decision making problem, AHP first extracts the hierarchy structure of the problem, and then the local weights of alternatives under each criterion are obtained by the eigenvector method. These weights are normalized as the sum of weights is equal to one. Finally, the global weights of alternatives are calculated by weighted additive sum. The AHP has been applied to various fields such as operations research, regional science and so on, since it is easier to use than other decision making tools, and since it can deal with the data difficult to be quantified.

However, AHP has been criticized for rank reversal phenomenon [2]. This phenomenon is that the ranks of alternatives can be reversed when a new alternative is added or an existing alternative is removed. Since the phenomenon invades the independence of alternatives, which is a major AHP assumption, it has been considered to be a contradiction of the methodology and some revisions have been proposed [3–6].

Recently, Triantaphyllou [7] reported that AHP sometimes made irrational rankings, of which the ranking obtained by AHP applied to all alternatives is different from the one obtained by the method to all pairs of alternatives, even when the transitivity property holds. He concluded in his paper that the phenomenon was caused by the additive sum to obtain the global weights, which is the last step of the AHP procedure, and he also presented that the multiplicative AHP (M-AHP) did not lead to the phenomenon.

This paper shows that the normalization method may also cause irrational rankings. In section 2, we first define the irrational ranking, and then give an example that the original AHP leads to the irrational rankings even when all pairwise comparisons are completely consistent. In section 3, we represent the cause of the irrational ranking and show that the Aspiration level AHP[1] (AL-AHP) [8,9], a revised AHP model proposed by Tamura et al., never lead to irrational rankings when all pairwise comparisons are completely consistent. We evaluate in section 4 how often irrational rankings occur for some AHP models by computational experiment in practical case, in which the all consistency indices (CI) of pairwise comparison matrices are less than 0.1. Finally in section 5, we give conclusions and subjects of future works.

2 Irrational Ranking

We first give the definition of the transitivity property.

Definition 1 (Transitivity property) For given alternatives A_i, $(i = 1, \cdots, m)$ and the binomial relation ($\preceq$), the transitivity property of the preference is said to hold if the condition

$$A_i \preceq A_j, A_j \preceq A_k \Rightarrow A_i \preceq A_k$$

is satisfied for any three alternatives A_i, A_j, A_k $(i, j, k \in \{1, \cdots, m\})$.

According to the definition, the alternatives are ranked uniquely if the transitivity property holds and if there is no indifferent alternative. On the other hand, when the transitivity property does not hold such as making a toss (that is, rock $\succ$ scissor, scissor $\succ$ paper and paper $\succ$ rock), the ranks of three choice (scissor, paper, rock) cannot be determined uniquely.

However, we sometimes observe different rankings by evaluating all alternatives simultaneously by AHP and those obtained from binomial relations of alternatives, in which transitivity holds, evaluated by AHP. In this paper, we call the phenomenon "irrational ranking." The next example illustrates irrational ranking.

Example 1 We evaluate three alternatives A_1, A_2, A_3 under three criteria C_1, C_2, C_3, whose weights are 4/22, 9/22, 9/22, respectively. We assume that all pairwise comparisons are completely consistent. The local and global weights obtained by AHP are shown in Table 1. From the table, we have the preference $A_3 \succ A_2 \succ A_1$.

Next, we evaluate binomial relations by AHP. The results are shown in Table 2. When pairwise comparison is completely consistent, the ratio of local weights is invariable. For example, the ratio of local weights in A_1 and A_2 under C_1 is 9 and 1 in Table 1 and is also same in Table 2. We have from Table 2 that the preference $A_1 \prec A_2$, $A_1 \prec A_3$ and $A_2 \succ A_3$. These results satisfy the transitivity property and we have $A_2 \succ A_3 \succ A_1$, which is different from the result obtained by evaluating all alternatives simultaneously.

[1] Aspiration level AHP is called Descriptive AHP and Behavioral AHP in [8] and [9], respectively.

Table 1. Weights obtained by evaluating all alternatives simultaneously

	$C_1(4/22)$	$C_2(9/22)$	$C_3(9/22)$	Weight
A_1	9/18	5/15	2/15	0.283
A_2	1/18	8/15	5/15	0.354
A_3	8/18	2/15	8/15	0.363

Table 2. Weights obtained by evaluating all pairs of alternatives

	$C_1(4/22)$	$C_2(9/22)$	$C_3(9/22)$	Weight
A_1	9/10	5/13	2/7	0.44
A_2	1/10	8/13	5/7	0.56
A_1	9/17	5/7	1/5	0.47
A_3	8/17	2/7	4/5	0.53
A_2	1/9	4/5	5/13	0.505
A_3	8/9	1/5	8/13	0.495

3 Cause and Several Revisions

We assume that w_j denotes the weight of the criterion C_j, $(j = 1, \ldots, n)$ and a_{ij} the i-th element of the principal eigenvector of the pairwise comparison matrix under the criterion C_j. a_{ij} is regarded as the weight of the alternative $i\,(i = 1, \ldots, m)$ before normalized. The total weight of the alternative k when evaluating all alternatives simultaneously with the original AHP can be written as (1). While the total weight of the alternative k when we evaluate two alternatives k and l with the original AHP is written as (2).

$$A_k = \sum_{j=1}^{n} w_j \frac{a_{kj}}{\sum_{i=1}^{m} a_{ij}} \tag{1}$$

$$A_k^{'} = \sum_{j=1}^{n} w_j \frac{a_{kj}}{a_{kj} + a_{lj}} \tag{2}$$

By comparing these equations, there may be the case of which the both $A_k > A_l$ and $A_k^{'} < A_l^{'}$ holds (See Example 1). This means the original AHP can lead to irrational rankings.

Triantaphyllou [7] presented that the irrational ranking was caused by the additive sum, which is the most popular way to integrate global weights in AHP. He also showed that the Multiplicative AHP (M-AHP) did not lead to irrational rankings. In M-AHP, the global weight of the alternative k is calculated by

$$\bar{A}_k = \prod_{j=1}^{n} \left(\frac{a_{kj}}{\sum_{i=1}^{m} a_{ij}} \right)^{w_j} . \tag{3}$$

Thus irrational ranking does not occur, since $\bar{A}_k > \bar{A}_l$ is equivalent to $\prod_{j=1}^{n}(a_{kj})^{w_j} > \prod_{j=1}^{n}(a_{lj})^{w_j}$, which means the weight ratio between $\bar{A}_k$ and $\bar{A}_l$ does not depend on the weights of the other alternatives.

By comparing the equations (1) and (3), we see that the additive sum does not cause irrational rankings if we may use the weight normalization method, in which the weight of an alternative does not depend on the other alternatives. AL-AHP [8,9], which was proposed by Tamura et al., is one of such methods. This revised AHP requires an aspiration level, a hypothetical alternative, under each criterion and the aspiration level is added to the set of alternatives to perform pairwise comparisons. Weights of alternatives are determined by the principal eigenvector of the pairwise comparison matrix normalized as the element corresponding to the aspiration level equal to one. The total weight of the alternative is calculated by the additive sum, which is same as the original AHP. For a detailed description of the procedure of AL-AHP, please refer to [8,9].

Theorem 1 AL-AHP does not cause the irrational ranking when pairwise comparisons are completely consistent.

Proof Let a_{0j} denote the weight of the aspiration level under the criterion C_j before normalized. Then the total weight of the alternative k is calculated by

$$\hat{A}_k = \sum_{j=1}^{n} w_j \frac{a_{kj}}{a_{0j}}, \tag{4}$$

when evaluating all alternatives simultaneously with AL-AHP. This is also the total weight of the alternative k when evaluating the two alternatives k and l with AL-AHP unless aspiration levels change. Therefore, the irrational ranking does not occur when using AL-AHP.

4 Experimental Analysis

We have already shown that the original AHP leads to irrational rankings even when all pairwise comparisons are completely consistent, and the both M-AHP and AL-AHP does not.

However, pairwise comparison is not always completely consistent in practical case. Saaty [1] introduced the consistency index (CI) of the pairwise comparison matrix, which represents how consistent the pairwise comparison is. CI is defined by

$$CI = \frac{\lambda_{\max} - n}{n - 1},$$

where n denotes the dimension of the matrix and $\lambda_{\max}$ denotes the principal eigenvalue of the matrix. CI is equal to zero when the pairwise comparison is completely consistent, and it becomes larger when the pairwise comparison is more inconsistent. Saaty [1] regards that the pairwise comparison is practically consistent when CI is less than 0.1 or 0.15.

We examine several AHP models for how often irrational ranking occurs under the condition that $CI \leq 0.1$ by computational experiment. In the experiment we evaluate three alternatives (A_1, A_2, A_3) under three criteria, and we assume that the pairwise comparison value is chosen from seven values, $\{1/7, 1/5, 1/3, 1, 3, 5, 7\}$.

We compare four models in the experiment. These are the original AHP, M-AHP which uses the multiplicative integration, AL-AHP which uses the aspiration level based normalization, and MAL-AHP which uses both the multiplicative integration and the aspiration level based normalization.

At first, we make all the 3-dimensional and 4-dimensional pairwise comparison matrices whose each element is one of the above seven pairwise comparison values, and remove the matrices not satisfying $CI \leq 0.1$. As a result, 121 matrices out of the total $7^3 = 343$ remain in the 3-dimensions and 6547 matrices out of the total $7^6 = 117649$ remain in the 4-dimensions. We next examine whether the transitivity property holds for the case where comparison matrices are selected from the remaining matrices. When the transitivity property holds, we also compare the ranking led from the binomial relations with the ranking obtained by evaluating all alternatives simultaneously. These comparisons are performed for all combinations of matrices for each four AHPs.

The results are shown in Table 3. The table shows that the transitivity property does not hold in about 1% of the cases in all AHPs. It also shows that irrational ranking occurs about 43.47% in the original AHP and 39.52% in the M-AHP. But, in AL-AHP and MAL-AHP, the ratio of irrational rankings decreases to less than 14%.

Table 3. Computational Results

	AHP	M-AHP	AL-AHP	MAL-AHP
Total cases	214,358,881	214,358,881	33,955,676,948,083	33,955,676,948,083
Anti-Transitivity	2,133,230	2,054,972	413,386,972,384	318,649,995,742
1st rank reversal	45,690,574	44,253,767	2,641,563,197,250	1,984,855,287,550
2nd/3rd rank reversal	46,565,199	39,643,403	1,977,030,291,391	1,708,868,121,307
Irrational ranking	92,255,773	83,897,170	4,618,593,488,641	3,693,723,408,857
Irrational Ratio	43.47%	39.52%	13.77%	10.98%

From these results, we can conclude that Saaty's consistency index is valid for preserving the transitivity property, but not appropriate for legal rankings. We can also conclude that both the weight normalization method with the aspiration level and the weight integration method with multiplicative form decrease the ratio of irrational rankings, but the normalization has much more effective to rid irrational rankings than the multiplicative integration.

5 Conclusion

We showed in this paper that the irrational ranking was caused by not only the additive sum for the global weight integration, but also the weight normalization method in which the sum of the principal eigenvector of the pairwise comparison matrix is equal to one. We also showed mathematically that M-AHP and AL-AHP never lead to such irrational rankings under the completely consistent conditions, and examined by computational experiment how often these models cause irrational rankings in practical case, in which CI of the pairwise comparison matrix is less than 0.1. The computational results showed that the weight normalization methods influence on irrational rankings greater than the weight integration methods. Thus we conclude that it may be desirable for rational decision making in AHP to introduce a certain standard alternative. Our future works are to propose a new criterion for examining the validity of the pairwise comparison matrices instead of Saaty's consistency index, and to propose a new AHP model which never causes irrational rankings when the pairwise comparisons are regarded to be valid.

References

1. Saaty T.L. (1990) The Analytic Hierarchy Process, 2nd ed. RWS Publications, Pittsburgh
2. Belton V., Gear T. (1983) On a Short-coming of Saaty's Method of Analytic Hierarchies, OMEGA The International Journal of Management Sciences 11, 3, 228–230
3. Belton V., Gear, T. (1985) The Legitimacy of Rank Reversal: A Comment, Omega 13, 3, 143–144
4. Shoner B., Wedley W.C. (1989) Ambiguous Criteria Weights in AHP: Consequences and Solutions, Decision Sciences 20, 3, 462–475
5. Shoner B., Wedley W.C., Choo, E.U. (1993) A unified approach to AHP with linking pins, European Journal of Operations Research 64, 384–392
6. Wedley W.C., Choo E.U., Schoner B. (1996) Benchmark Measurement: Between Relative and Absolute. In: Proceedings of the Fourth International Symposium on the Analytic Hierarchy Process. Simon Fraser University, B.C., Canada, July 12–15, 1996. 335–345
7. Triantphyllou E. (2001) Two New Cases of Rank Reversal when the AHP and Some of its Additive Variants are Used that do not Occur with Multiplicative AHP, Journal of Multi-Criteria Decision Analysis 10, 11–25
8. Tamura H., Takahashi S., Hatono I., Umano, M. (1998) On a Descriptive Analytic Hierarchy Process (D-AHP) for Modeling the Legitimacy of Rank Reversal (In Japanese), Journal of the Operation Research Society of Japan 41, 2, 214–227
9. Tamura H., Takahashi S., Hatono I., Umano, M. (2000) On a Behavioral Model of Analytic Hierarchy Process for Modeling the Legitimacy of Rank Reversal. In: Research and Practice in Multiple Criteria Decision Making. Springer, Berlin Heidelberg, 173–184

Choquet Integral Type DEA

Eiichiro Takahagi

SENSHU University, Kawasaki, Kanagawa, 214-8580, JAPAN

Abstract. Choquet integral type DEA is a global evaluation model of cross efficiency scores which calculated by ϕ_s transformation type fuzzy measure and Choquet integral model. As D-efficient DMU's input and output weights are not determined uniquely, the DMU's cross-efficiency scores are not determined uniquely. So, we propose the maximum model and the averaging model of the scores.

1 Introduction

CCR models evaluate DMUs (decision making units) using most favorable input and output weights for the DMU, i.e. most favorable evaluation. The weights are the direction that the DMU aims, that is the DMU's evaluation method. Therefore, this evaluation is one of self-evaluations. Cross-efficiency scores are evaluation scores that are evaluated by other DMU's input and output weights. For one DMU, we get n cross-efficiency scores (n is the number of DMUs). The averaging scores are global evaluations of DMUs. If we use maximum values of cross-efficiency scores (one of averaging score), the evaluation is most optimistic evaluation, that is CCR model's efficiency score.

In this paper, we do intermediate evaluations among minimum, average and maximum evaluations by varying parameter ξ. Using line charts, we enable graphical representations of DMU's global evaluations.

2 Fuzzy Measure Choquet Integral Model

Fuzzy measure Choquet integral models are averaging models that enable averaging evaluation among maximum and minimum including weighted arithmetic average[4].

Let $X = \{1, \ldots, n\}$ be a set of inputs of Choquet integrals and μ be a fuzzy measure satisfying $\mu : 2^X \rightarrow [0,1], \mu(\emptyset) = 0, \mu(X) = 1$ and $A \subseteq B \subseteq X \rightarrow \mu(A) \leq \mu(B)$.

Let $h(1), \ldots, h(n), h(i) \in [0,1]$ be input variables. Choquet integral is defined as (C) $\int h d\mu \equiv \int_0^1 \mu(\{x; h(x) > r\})dr$.

The averaging evaluation is calculated by ϕ_s transformation[1] from every inputs' weights and interaction index ξ such that:

$$\phi_s(p) = \begin{cases} \langle p \rangle & \text{if } s = 0 \\ p & \text{if } s = 1 \\ 1 - \langle 1 - p \rangle & \text{if } s = +\infty \\ (s^p - 1)/(s - 1) & \text{otherwise} \end{cases}, \text{where } \langle p \rangle = \begin{cases} 1 & \text{if } 0 < p \leq 1 \\ 0 & \text{if } p = 0 \end{cases}$$

and $s = (1-\xi)^2/\xi^2$ (if $\xi = 0$ then $s = +\infty$) Fig. 1 shows ξs meaning. A fuzzy measure μ_ξ is assigned from ξ and each input weights ($w_i, w_i \in [0,1], i = 1,\ldots,n, \sum w_i = 1$):

$$\mu_\xi(A) = \phi_s(\sum_{i \in A} w_i). \tag{1}$$

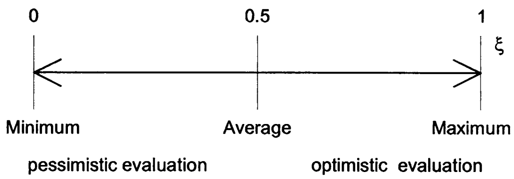


Fig. 1. ξ's means

3 CCR Model (Notations)

DMU_o :o-th DMU$o = 1, 2, \ldots, n$
x_{io} :DMU_o's i-th input value (nonnegative),$i = 1, \ldots, m$
y_{jo} :DMU_o's j-th output value (nonnegative),$j = 1, \ldots, k$
v_{io} :DMU_o's i-th input weight
u_{jo} :DMU_o's j-th output weight
θ_{ao} : DMU_a 's cross-efficiency score using DMU_o's weights

We get DMU_o's efficiency score θ_{oo} by following LP;

Maximize: $\theta_{oo} = (\sum_{j=1}^{k} u_{jo} y_{jo})/(\sum_{i=1}^{m} v_{io} x_{io})$ (2)

Subject to: $(\sum_{j=1}^{k} u_{jo} y_{ja})/(\sum_{i=1}^{m} v_{io} x_{ia}) \leq 1, a = 1, \ldots, n, v_{io}, u_{jo} \geq 0, \forall i, j$,

and cross-efficiency scores θao by

$$\theta_{ao} = (\sum_{j=1}^{k} u_{jo} y_{ja})/(\sum_{i=1}^{m} v_{io} x_{ia}). \tag{3}$$

4 Choquet Integral Type DEA (Maximum Model)

4.1 Global Evaluations

Using Choquet integral models, we can do intermediate evaluation among minimum, average and maximum varying ξ. From DMU_a's cross-efficiency scores $\theta_{a1}, \ldots, \theta_{an}$, we can get the DMU_a's global evaluations $\theta^*_{\xi,a}$.

$$\theta^*_{\xi,a} = (\mathrm{C}) \int h^a d\mu_\xi, \; h^a(i) = \theta_{ai}, i = 1, \ldots, n \tag{4}$$

If we use $\xi = 1$, this model is the same as CCR model, that is $\theta_{aa} = \theta^*_{1,a}, \forall a$.

As the μ_ξ is calculated by (1), we must identify ξ and w_i. ξ is identified by Fig.1's means. To do sensitivity analysis, we vary the ξ in $[0, 1]$. In CCR models, i.e. $\xi = 1$, the w_is are ignored, because, CCR model are maximum evaluations of cross-efficiency scores. In this model, we can introduce the weights, that mean importance or impacts of DMUs. Default values are $w_i = 1/n$.

4.2 Uniqueness of Efficiency DMU's Weights

If DMU_o is a D-efficient DMU, that is $\theta_{oo} = 1$, u_{jo} and $v_{jo} (j = 1, \ldots n)$ are not determined uniquely. So, cross-efficiency scores $\theta_{ao} (\forall a \in \{1, \ldots, n\} \setminus \{o\})$ are not determined uniquely. Sexton[2], Hibiki[3] and many researchers proposed some solutions. In this section, we use maximum adjusted cross-efficiency scores. In this method, we make sub DEA model S that is maximize own efficiency value under the $\theta^S_o = 1$, where θ^S_o is the DMU_o's cross-efficiency score of the sub model:

$$\begin{aligned} \text{Maximize:} \quad & (\theta_{ao} =)\theta^S_{aa} = (\textstyle\sum_{j=1}^k u^S_{ja} y_{ja})/(\sum_{i=1}^m v^S_{ia} x_{ia}) \qquad (5) \\ \text{subject to:} \quad & (\textstyle\sum_{j=1}^k u^S_{ja} y_{jb})/(\sum_{i=1}^m v^S_{ia} x_{ib}) \leq 1, \forall b \in \{1, \ldots, n\} \setminus \{o\} \\ & \theta^S_o = (\textstyle\sum_{j=1}^s u^S_{ja} y_{jo})/(\sum_{i=1}^m v^S_{ia} x_{io}) = 1 \,,\, v^S_{ia}, u^S_{ja} \geq 0, \forall i, j, a, \end{aligned}$$

where u^S_{ja} and v^S_{ia} are sub model's output and input weights.

5 Choquet Integral Type DEA(Average Model)

In §4, sub model's global evaluation scores are maximum values of cross-efficiency scores. However, by varying ξ, we can do many global evaluations. So, from sub model's cross-efficiency scores, we calculate global evaluation scores with original model's ξ.

In original model, if DMU_o is D-efficient, we make sub model S with $\theta_o = 1$. From the LP model (5), we can get sub model's input and output weights $(v^S_{ia}, u^S_{ja} \geq 0, \forall i, j, a)$. From those weights, we can calculate cross-efficiency scores:

$$\theta^S_{ab} = (\sum_{j=1}^{s} u^S_{jb} y_{ja})/(\sum_{i=1}^{m} v^S_{ib} y_{ia}), \forall a \in \{1, \ldots, n\} \setminus \{o\}. \qquad (6)$$

Using the θ^S_{ab}, we calculate global evaluation θ_{ao}, such that:

$$\theta_{ao} = \theta^{S*}_{\xi o} = (\mathrm{C}) \int h^a d\mu_\xi, \; h^a(b) = \theta^S_{ab}, b = 1, \ldots, o-1, o+1, \ldots, n. \quad (7)$$

However, in the sub model S, it is possible that there are D-efficient DMUs. Let a D-efficient DMU in the sub model S be p. So u^S_{jp} and $v^S_{jp} (\forall j)$ are not determined uniquely. Then, cross-efficiency scores θ^S_{ap} are neither determined. In this case, we make sub model T, that is a DEA model under the $\theta^T_o = 1$ and $\theta^T_p = 1$:

$$\begin{aligned} \text{Maximize :} \quad & \theta^T_{aa} = (\textstyle\sum_{j=1}^k u^T_{ja} y_{ja})/(\textstyle\sum_{i=1}^m v^T_{ia} x_{ia}) \qquad (8) \\ \text{subject to:} \quad & (\textstyle\sum_{j=1}^k u^T_{ja} y_{jb})/(\textstyle\sum_{i=1}^m v^T_{ia} x_{ib}) \le 1 \,,\, \forall b \in \{1,\ldots,n\} \setminus \{o,p\} \\ & \theta^T_o = (\textstyle\sum_{j=1}^s u^T_{ja} y_{jo})/(\textstyle\sum_{i=1}^m v^T_{ia} x_{io}) = 1 \\ & \theta^T_p = (\textstyle\sum_{j=1}^s u^T_{ja} y_{jp})/(\textstyle\sum_{i=1}^m v^T_{ia} x_{ip}) = 1 \,,\, v^T_{ia}, u^T_{ja} \ge 0, \forall i,j,a \end{aligned}$$

If there is D-efficient DMU_q in sub model T, we make sub model U under the $\theta^U_s = \theta^U_p = \theta^U_q = 1$. This process continues until not existing D-efficient DMUs.

Table 1. Input-Output Data

DMU	x_1	x_2	x_3	x_4	x_5	y_1	y_2
D1	2.249	163.523	26	158.713	49.196	5.561	105.321
D2	4.617	338.671	30	73.756	78.599	18.106	314.682
D3	3.873	281.655	51	149.881	176.381	16.498	542.349
D4	5.541	400.993	78	166.155	189.397	30.810	847.872
D5	11.381	363.116	69	311.548	192.235	57.279	758.704
D6	10.086	541.658	114	379.632	194.091	66.137	1438.746
D7	5.434	508.141	61	176.388	228.535	35.295	839.597
D8	7.524	338.804	74	203.489	238.691	33.188	540.821
D9	5.077	511.467	84	210.652	267.385	65.391	1562.274
D10	7.029	393.815	68	251.715	277.402	41.197	978.117
D11	11.121	509.682	96	308.207	330.609	47.032	930.437

6 Numerical Examples

6.1 Maximum Model

Table 1 is the input-output data of 11 DMUs. Table 2 and Fig.2 is the result of maximum model.

- As D9 is always 1, D9 get always D-efficient evaluations from all DMUs .
- D5 and D6 are D-efficient DMUs, but evaluations of D5 and D6 in view of other DMUs are low. So, in Fig.2, D5 and D6's decrease greatly and almost DMUs do not support the D5 and D6's good self-evaluations.
- In Fig.2, D4's line is almost horizon, so D4 is not dependent on other DMU's evaluations.

Table 2. Maximum model

DMU	Cross-Efficiency Scores											Aggregation		
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	MIN	Average	MAX
D1	0.35	0.25	0.21	0.23	0.35	0.35	0.27	0.21	0.35	0.24	0.24	0.21	0.28	0.35
D2	0.76	0.90	0.30	0.66	0.89	0.90	0.75	0.47	0.90	0.37	0.37	0.30	0.66	0.90
D3	0.39	0.37	0.63	0.52	0.53	0.58	0.44	0.43	0.63	0.53	0.53	0.37	0.51	0.63
D4	0.59	0.65	0.69	0.75	0.66	0.75	0.52	0.60	0.75	0.64	0.64	0.52	0.66	0.75
D5	1.00	0.95	0.68	0.57	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.57	0.93	1.00
D6	1.00	1.00	0.87	1.00	1.00	1.00	0.74	0.83	1.00	0.92	0.92	0.74	0.93	1.00
D7	0.65	0.63	0.54	0.63	0.74	0.67	0.74	0.56	0.74	0.54	0.54	0.54	0.64	0.74
D8	0.54	0.56	0.52	0.38	0.70	0.61	0.55	0.70	0.70	0.66	0.66	0.38	0.60	0.70
D9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
D10	0.63	0.58	0.81	0.59	0.82	0.75	0.78	0.73	0.82	0.82	0.82	0.58	0.74	0.82
D11	0.56	0.55	0.60	0.47	0.67	0.63	0.61	0.66	0.67	0.67	0.67	0.47	0.61	0.67

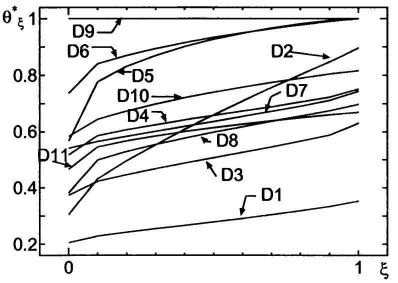


Fig. 2. Maximum Model

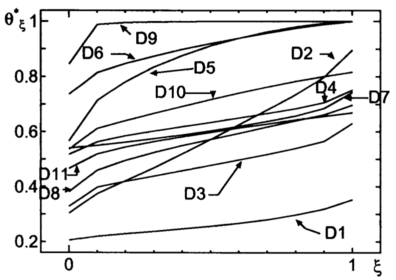


Fig. 3. Average Model

6.2 Averaging Model

Fig.4 is the procedure of averaging model.

1. We calculate the original model, but D5,D6,D8 and D9 are D-efficient DMU.
2. To calculate the D5's cross-efficiency scores, we make sub model S, but D6 and D9 are D-efficient DMU in S.
3. To calculate the D6's cross-efficiency score of sub model S, we make sub model T, but D9 is D-efficient DMU in T.
4. To calculate the D9's cross-efficiency score of sub model T, we make sub model U. In sub model U, all DMUs are not D-Efficient, all cross-efficiency scores are determined uniquely. So, from the cross-efficiency scores, using original model's ξ , we can calculate global evaluation scores of U, that are D9's cross-efficiency score of sub model T.
5. In this way, all cross-efficiency scores are calculated from each sub model's global evaluations.

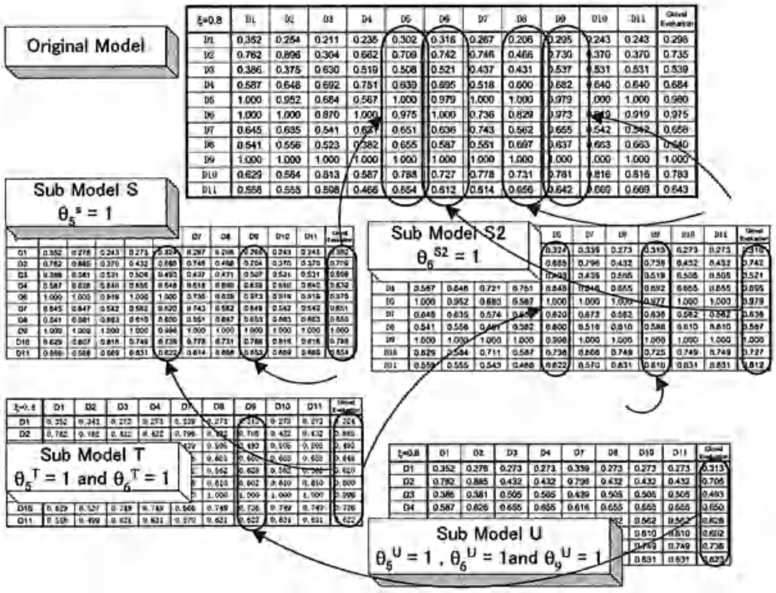


Fig. 4. Procedure of Averaging Model

7 Conclusions

We present global evaluation methods of cross efficiency scores that enable pessimistic and optimistic view by varying parameter ξ. The outputs of our model are line charts like Fig.2 and 3. Comparing the two charts,

- in maximum model, D9 is always 1, but in averaging model, not
- other DMU's rows are similar variations.

So, without strictly analyses, we proper to use maximum model, because averaging model's calculation amount is big.

References

1. Tukamoto Y. (1982) A Measure Theoretic Approach to Evaluation of Fuzzy Set Defined on Probability Space, Journal of Fuzzy Math , Vol. 2, No.3 89-98
2. Sexton T. F. et al. (1986) Data Envelopment Analysis: Critique and Extensions, in R.H. Silkman (ed.) Measuring Efficiency: An Assessment of Data Envelopment Analysis, 73-105.
3. Hibiki N. (1995) On Uniquely Identification Methods of DEA's Adjusted Cross-efficiency Scores, Technical Report No.95004, Department of Administration Engineering, Factory of Science and Technology, Keio University (Japanese).
4. Takahagi E. (2000) On Identification methods of λ-fuzzy measure using weights and λ, Journal of Japan Society for Fuzzy Theory and Systems, Vol. 12, No.5 665-676 (Japanese).

Interactive Procedures in Hierarchical Dynamic Goal Programming*

T. TRZASKALIK

Department of Operations Research,

The Karol Adamiecki University of Economics,

ul. Bogucicka 14, 40-226 Katowice, Poland,

ttrzaska@ae.katowice.pl

Abstract. The paper is devoted to application of period target values in hierarchical goal dynamic programming. Period backward approach is considered. An interactive procedure is proposed an numerical example is also provided. The presented approach can be extended to multiperiod target values, changeable hierarchy of criteria and forward separable vector criteria functions. It can be found in Trzaskalik (1998).

Keywords. Dynamic programming, vector optimization, goal programming, interactive procedure.

1. Discrete Multi-Objective Dynamic Programming Problem

We consider a discrete decision process that consists of T periods. Let Y_t be the set of all feasible state variables for period t and $X_t(y_t)$ – the set of all feasible decision variables for period t and state $y_t \in Y_t$ (t=1,...,T). Set Y_{T+1} includes all the states at the end of the process. We assume that all these sets are finite. Let $y_t \in Y_t$ and $x_t \in X_t(y_t)$. Period realization is defined as $d_t \cong (y_t, x_t)$. D_t is the set of all period realizations in period t. We assume that for t=1,...,T there are given transformations $?_t$: D_t ? Y_{t+1}. We denote d as a process realization. Set of all process realizations is defined as follows:

$$D \cong \{d=(d_1,...,d_T): \forall_{t\in 1,...,T} \ \ y_{t+1} = ?_t(y_t, x_t), \ \ x_t \in X_t(y_t) \}.$$

We assume, that in each period t there are defined K period criteria functions F_t^k: D_t? **R**, (k=1,...,K, t=1,..., T).

Each M-dimensional (M=2) function with components $F^m \cong F^m(F_1^1, F_1^2,...,F_T^K)$ for m=1,...,M where F^m is a scalar function can be considered as vector-valued criterion function. Let $F \cong [F^1,...,F^M]'$. Components F^m are

* Research partly supported by the 2H02B00622 grant of the Polish State Committee for Scientific Research

called multiperiod criteria functions. In the further considerations we postulate maximization of all components of F.

Let us assume that there are given two process realizations d, d' and vectors $F(d) \cong [F^1(d),...,F^m(d)]'$ and $F(d') \cong [F^1(d'),...,F^m(d')]'$. Relation of domination = is defined as follows:

$$F(d) = F(d') \Leftrightarrow \forall_{m=1,...,M} F^m(d) = F^m(d) \wedge \exists_{i=1,...,M} F^i(d) > F^i(d')$$

If $F(d) = F(d')$, vector F(d) dominates vector F(d') and realization d is better than realization d'. Realization d^* is efficient, if $\sim\exists_{d\in D} F(d) = F(d^*)$. $\overset{*}{D}$ is the set of all efficient realizations. Set

$$\overset{*}{D}(\bar{d}) \cong \{d^* \in \overset{*}{D} : F(d^*) \geq F(\bar{d})\}$$

consists of all efficient realizations which are better than $\bar{d}$.

F^m is scalar backward separable, if there exist functions $f_t^m(F_t^1,...,F_t^K)$ and operators $\circ_t^m$ (m=1, ...,M; t=1,...,T-1) such that

$$F^m = f_1^m \circ_1^m (f_2^m \circ_2^m (... (f_{T-1}^m \circ_{T-1}^m f_T^m)...))$$

F is backward separable, if each component F^m is scalar backward separable. In a similar way we can formulate the definition of forward separability [Trzaskalik (1998)]. If F is backward separable and backward monotone (or forward separable and forward monotone) we can apply Bellman's principle of optimality to solve dynamic vector optimization problem [Belmann (1957), Trzaskalik (1998), Li and Haimes (1989)]. We can also find sets $\overset{*}{D}(\bar{d})$ for each $\bar{d} \in D$.

2. Goal Programming Approach

Let Γ^m be period matrix containing values of period goals

$$G^m \cong \begin{bmatrix} ?_1^1 & ?_2^1 & \cdots & ?_T^1 \\ \cdots & \cdots & \cdots & \cdots \\ ?_1^M & ?_2^M & \cdots & ?_T^M \end{bmatrix}$$

and let $\overset{+}{C}, \overset{-}{C}$ be forward weight coefficients matrices containing penalty coefficients for multiperiod deviations from target goals for $m \in 1,...,M$, $t \in 1,...T$:

$$\overset{+}{C} \cong \begin{bmatrix} \overset{+1}{c_1} & \overset{+1}{c_2} & \dots & \overset{+1}{c_T} \\ \dots & \dots & \dots & \dots \\ \overset{+M}{c_1} & \overset{+M}{c_2} & \dots & \overset{+M}{c_T} \end{bmatrix} \quad \overset{-}{C} \cong \begin{bmatrix} \overset{-1}{c_1} & \overset{-1}{c_2} & \dots & \overset{-1}{c_T} \\ \dots & \dots & \dots & \dots \\ \overset{-M}{c_1} & \overset{-M}{c_2} & \dots & \overset{-M}{c_T} \end{bmatrix}$$

We define auxiliary functions

$$\overset{+}{h}{}_t^m \cong \begin{cases} f_t^m(d_t) - ?_t^m & \text{if } f_t^m(d_t) - ?_t^m \geq 0 \\ 0 & \text{if } f_t^m(d_t) - ?_t^m < 0 \end{cases}$$

$$\overset{-}{h}{}_t^m \cong \begin{cases} 0 & \text{if } f_t^m(d_t) - ?_t^m \geq 0 \\ ?_t^m - f_t^m(d_t) & \text{if } f_t^m(d_t) - ?_t^m < 0 \end{cases}$$

Let

$$s_t^m(d_t) \cong \overset{+}{c}{}_t^m \overset{+}{h}{}_t^m(d_t) + \overset{-}{c}{}_t^m \overset{-}{h}{}_t^m(d_t)$$

and

$$\overset{\Leftarrow}{s}{}^m(d) \cong \sigma_1^m(d_1) \circ_1^m (\sigma_2^m(d_2) \circ_2^m (\dots(\sigma_{T-1}^m(d_{T-1}) \circ_{T-1}^m \sigma_T^m(d_T)\dots)))$$

Dynamic goal programming problem can be formulated as follows

$$\text{Min}\ \{\sum_{m=1}^{M} \overset{\Leftarrow}{s}{}^m(d) : d \in D\}.$$

3. Hierarchical Goal Programming Approach

We assume that a backward process is given and we deal with fixed single hierarchy of criteria. It means that multi-period criteria are numerated in such a way that a more important one has the lower number then any less important one. The position of each criterion is the same in all the periods. Procedure submitted below gives the decision maker a possibility of interactive modeling period backward fixed single hierarchy target goal structure of the final solution.

Procedure (period backward approach)

1. Let $m \cong 1$ and $D^0 \cong D$.

2. The DM sets period target goal vectors

$$\gamma^m \cong [\,?_1^m, \dots, ?_T^m\,].$$

and period tolerance limit vectors

$$\overset{+}{?}{}^{m} \cong [\overset{+}{?}{}^{m}_{1}, \ldots, \overset{+}{?}{}^{m}_{T}],\ \overset{-}{?}{}^{m} \cong [\overset{-}{?}{}^{m}_{1}, \ldots, \overset{-}{?}{}^{m}_{T}], \qquad \overset{+}{?}{}^{m}_{t}, \overset{-}{?}{}^{m}_{t} \geq 0.$$

3. Create relaxed sets:

$$D^{m} \cong \{d \in D^{m-1} : \forall_{t \in 1,\ldots,T} \ \overset{+}{h}{}^{m}_{t}(d_t) \leq \overset{+}{?}{}^{m}_{t} \wedge \overset{-}{h}{}^{m}_{t}(d_t) \leq \overset{-}{?}{}^{m}_{t} \}$$

If necessary, apply the procedure of generating feasible states and decisions [Trzaskalik (1998)].

4. If m≤M-1, set m ≅ m+1 and proceed to step 2.

5. Applying the backward single criterion dynamic programming approach, solve problem:

$$\text{Min}\ \{ \overset{\Leftarrow}{s}{}^{M}(d) : d \in D^{m} \}.$$

6. Let $\bar{D}$ be the set of solutions of the problem solved in step 5. The DM chooses realization $\bar{d} \in \bar{D}$.

7. Set $\overset{*}{D}(\bar{d})$ is created.

8. The DM makes one of the following decisions:

a) Accepts $\bar{d}$ as the final realization,

b) Chooses one of realizations from $\overset{*}{D}(\bar{d})$ as the final one,

c) Points to another solution from $\bar{D}$ and goes back to step 8,

d) Repeats the whole procedure with changed target goal vectors or changed parameters.

e) Breaks the procedure.

4. Numerical Example

Assume that the DM defines the following hierarchy of criteria: the most important is criterion 1, the less important – criterion 2. The DM sets a target goal vector and period tolerance limit vectors for criterion 1 as follows:

$$\gamma^{1} = [5, 7, 9,]; \quad \overset{+}{?}{}^{1} = [2, 2, 0]; \quad \overset{-}{?}{}^{1} = [2, 2, 6].$$

The graph of the process with values $\overset{+}{h}{}^{1}_{t}(y_t, x_t)$, $\overset{-}{h}{}^{1}_{t}(y_t, x_t)$ is given in Fig. 1.

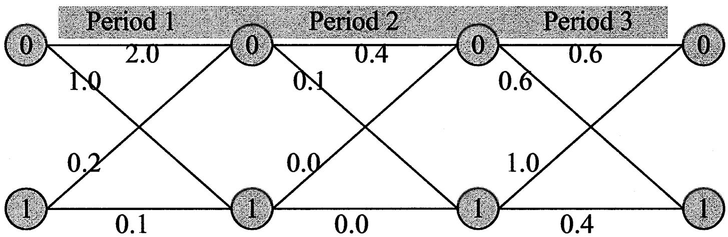


Fig. 1. The graph of the exemplary process.

For m=1 we consider such period realizations, for which

$$\overset{+}{h}{}^1_1(y_1, x_1) \le 2 \;\wedge\; \overset{-}{h}{}^1_1(y_1, x_1) \le 2,$$

$$\overset{+}{h}{}^1_2(y_1, x_1) \le 2 \;\wedge\; \overset{-}{h}{}^1_2(y_1, x_1) \le 2,$$

$$\overset{+}{h}{}^1_3(y_1, x_1) \le 0 \;\wedge\; \overset{-}{h}{}^1_3(y_1, x_1) \le 6.$$

The graph of the process after the first reduction is shown in Fig. 2.

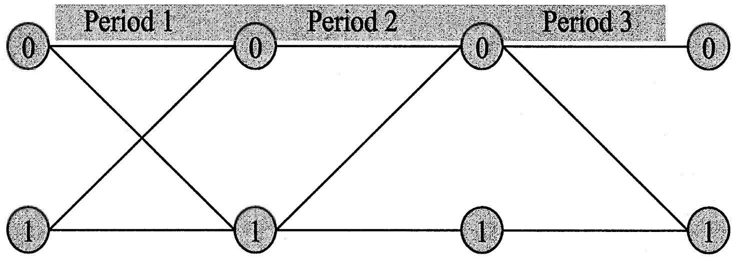


Fig. 2. The graph of the process after the first reduction

Assume that the DM sets target goal vector and period tolerance limit vectors for criterion 2 as follows:

$$\gamma^2 = [8, 7, 8]; \;\; \overset{+}{?}{}^1 = [1, 1, 1]; \;\; \overset{-}{?}{}^1 = [2, 2, 2].$$

The graph of the process with values $\overset{+}{h}{}^2_t(y_t, x_t)$, $\overset{-}{h}{}^2_t(y_t, x_t)$ is given in Fig. 3.

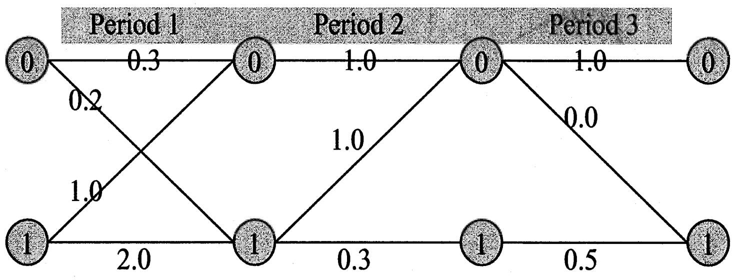


Fig. 3. The graph of the process after the first reduction with the additional values.

For m=2 we consider such period realizations, for which

$$\overset{+}{h}{}^2_1(y_1, x_1) \le 1 \;\wedge\; \overset{-}{h}{}^2_1(y_1, x_1) \le 2,$$

$$\overset{+}{h}{}_2^2(y_1, x_1) \le 1 \;\wedge\; \overset{-}{h}{}_2^2(y_1, x_1) \le 2,$$

$$\overset{+}{h}{}_3^2(y_1, x_1) \le 1 \;\wedge\; \overset{-}{h}{}_3^2(y_1, x_1) \le 2.$$

The graph of the process after the second reduction is shown in Fig. 4.

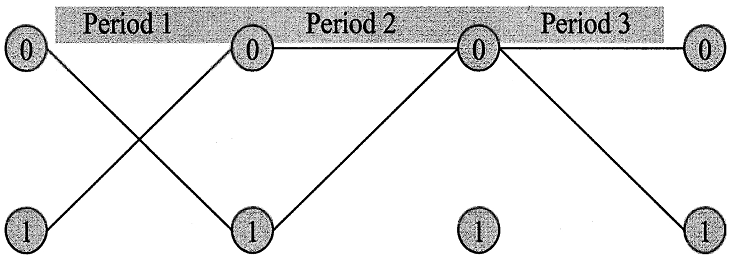


Fig. 4. The graph of the process after the second reduction

We have:

$$D^2 = \{\overset{1}{\bar{d}}, \overset{2}{\bar{d}}, \overset{3}{\bar{d}}\}$$

where $\overset{1}{\bar{d}}$=(0, 1, 1, 0, 0,1), $\overset{2}{\bar{d}}$=(0, 1, 1, 0, 0,1), $\overset{3}{\bar{d}}$=(0, 1, 1, 0, 0,1).

We obtain

$$\overset{\Leftarrow}{s}{}^M(\overset{1}{\bar{d}}) = 3, \quad \overset{\Leftarrow}{s}{}^M(\overset{2}{\bar{d}}) = 4, \quad \overset{\Leftarrow}{s}{}^M(\overset{3}{\bar{d}}) = 3.$$

We are to solve problem:

$$\text{Min}\,\{\overset{\Leftarrow}{s}{}^M(d) \colon d \in D^M\}$$

We obtain $\bar{D} = \{\overset{1}{\bar{d}}, \overset{3}{\bar{d}}\}$. The DM is asked to make the final choice between these realizations.

Literature:

Bellman R. (1957): Dynamic Programming, Princeton University Press.

Li D., Haimes Y. Y. (1989): Multiobjective dynamic programming: the state of the art. Control-Theory and Advanced Technology, 5, 4, 471-483.

Trzaskalik T. (1998): Multiobjective Analysis in Dynamic Environment, University of Economics in Katowice.

Solution Concepts for Coalitional Games in Constructing Networks

Masayo Tsurumi[1], Tetsuzo Tanino[2], and Masahiro Inuiguchi[2]

[1] Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan
[2] Department of Electronics and Information Systems, Graduate School of Engineering, Osaka University, 2-1 Yamada-oka, Suita, Osaka 565-0871, Japan

Abstract. In this paper, we shall discuss solution concepts for coalitional games in constructing networks. We define the demand allocation using a concept of the demand operation which was defined in our previous paper. A sufficient condition that demand allocations belong to the core of coalitional games in constructing networks has been shown. By using this theorem, we have obtained a sufficient condition for the core of coalitional games in constructing networks with public vertices to be nonempty.

1 Introduction

In recent years, to discuss optimization problems concerned with constructing networks such as information systems is getting more important[4,7–9]. In these optimization problem, players are willing to construct some distribution system with the minimum cost. Once a distribution system is built with the minimum cost, the problem of how to allocate the cost to each member will arise. Such a problem was first introduced by Claus and Kleitman [2].

Bird [1] was the first who suggested a game theoretic approach to this problem. He proposed rational allocations called Bird tree allocations. A Bird tree allocation is an element of the cores of any coalitional games concerned with constructing networks[3,5]. However, Bird tree allocations are not so realistic. It is because a player directly connected to the supplier does not receive benefit by forming the grand coalition, although his/her cooperation is important for constructing a link between the supplier and other players in many cases. To improve this point, Granot and Huberman [6] proposed weak demand operation (w.d.o.) by a player and weak demand operation by a coalition. However, the w.d.o. by a coalition is not well-defined. Furthermore, there exist some coalitional games whose cores do not include allocations obtained by weak demand operation.

In our previous paper [9], we have added a certain restriction to the weak demand operation by a coalition so that it is well-defined. However, some of the obtained allocations still do not belong to the core of all coalitional games for constructing networks. Hence, we have proposed a new concept of demand operation. Any allocation obtained through the operation on a Bird tree allocation is an element of the cores of any coalitional games concerned with constructing networks[9].

On the other hand, Granot et al. [4] dealt with coalitional games concerned with constructing networks with public vertices, which are not inhabited by any player and which are available for all players. Granot showed a sufficient condition for the core to be nonempty.

In this paper, we discuss coalitional games in constructing networks. We define a new concept of demand operation by a coalition, which will be called a demand allocation. A sufficient condition that demand allocations belong to the core of coalitional games in constructing networks is shown. By using this condition, we shall give a sufficient condition for the core of coalitional games in constructing networks with public vertices to be nonempty.

2 Games in Constructing Networks

Let $N = \{1, 2, \dots, n\}$ be a set of players. Then a function $c : 2^N \to \mathbb{R}_+$ satisfying $c(\emptyset) = 0$ is said to be a discrete cost function or a cost-sharing game (a game) if $c(S)$ can be regarded as the cost for a coalition S, where $\mathbb{R}_+ = \{r \in \mathbb{R} \mid r \geq 0\}$. Let x_i denote the amount charged to player i. An n-vector $\boldsymbol{x} = (x_1, \dots, x_n)$ satisfying $x(N) = c(N)$ is called an allocation. Then the core of a game c is defined as follows:

$$Core(c) = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \sum_{i \in N} x_i = c(N), \quad \sum_{i \in S} x_i \leq c(S), \forall S \subset N \right\}.$$

Let (V, E) denote an undirected graph with the set of nodes V and the set of vertices E. An edge $e \in E$ is denoted by (u, v) if the end points of e are u and v. Edges (u, v) and (v, u) denote the same edge for any $u, v \in V$. If there exists an edge (u, v) in a graph G for any $u, v \in V$, G is said to be a complete. Consider a sequence $P = (v_0, e_1, v_1, \dots, e_l, v_l)$, where $v_0, \dots, v_l \in V$ and $e_1, \dots, e_l \in E$. A sequence P is said to be a path if $e_i = (v_{i-1}, v_i)$ for any $i \in \{1, 2, \dots, l\}$. A path P is said to be a circuit if v_0 and v_l denote the same vertex. A graph G is said to be connected if there exists a path from u to v for any $u, v \in V$. It is clear that a graph G is connected if it is complete. A connected subgraph without any circuits is called a tree. A tree is said to be a spanning tree for V' if its vertex set is equal to V'.

In this paper, (V, E, w) is called a network if (V, E) is a graph and $w : E \to \mathbb{R}_+$. For $e = (u, v)$, we regard $w(e) = w_{uv} \geq 0$ as the cost of constructing the link between u and v. It is assumed that the underlying graph (V, E) in (V, E, w) is complete without loss of generality. For a network (V', E', w') with $w' : E' \to \mathbb{R}_+$, a spanning tree is said to be a minimum spanning tree for V' and denoted by $\Gamma[V']$ if it is minimum of all spanning trees for V' in terms of the cost. Note that some distinct minimum spanning trees exist in many networks. Let $(V_{\Gamma'}, E_{\Gamma'})$ denote a minimum spanning tree Γ'. If a vertex v_1 is on the unique path from the source $*$ to a vertex v_2 in a minimum spanning tree Γ' including a source $*$, the notation $v_1 \preceq_\Gamma v_2$ is used and vertices v_1 and v_2 are called a predecessor of v_2 and a follower of v_1, respectively. In particular, if $(v_1, v_2) \in E_\Gamma$ and $v_1 \preceq_\Gamma v_2$ then the vertices v_1 and v_2 is said to be the adjacent predecessor of vertex v_2 and an adjacent follower of vertex v_1 in Γ, respectively. Denote the adjacent predecessor of v_2 by $p(v_2)$ and the set of adjacent follower of v_1 by $F(v_1)$. For $V' \subseteq V$, let $F(V') = \cup_{v \in V'} F(v) \backslash V'$.

In this paper, $(V \cup \{*\}, E, w, N, a)$, where N is a player set and $a : N \to V$ is an injection, is called a spanning network problem (SN-problem). We may identify the vertex $a(i)$ with $i \in N$, if there is no fear of confusion. Let $A(S) = \cup_{i \in S}\{a(i)\}$ and $A_*(S) = A(S) \cup \{*\}$. A spanning tree which is minimum of all spanning trees including $A(N)$ in terms of cost is called minimum spanning trees w.r.t. $(V \cup \{*\}, E, w, N, a)$ and denoted by Γ if there is no fear of confusion. A vertex $v \in V \backslash A(N)$ is called a public vertex. The set of public vertices is denoted by P, i.e., $P = V \backslash A(N)$. Here, assume that public vertices are available for any player, then we have the following definition of the minimum spanning tree game (cf. [3,5]):

Definition 1. A function $\hat{c} : 2^N \to \mathbb{R}_+$ is said to be a minimum spanning tree games (mst-games) w.r.t. an SN-problem $(V \cup \{*\}, E, w, N, a)$ if the following holds.

$$\hat{c}(S) = \min_{A_*(S) \subseteq V' \subseteq A_*(S) \cup P} \sum_{e \in E_{\Gamma[V']}} w(e), \qquad \forall S \subset N,$$

where $\hat{c}(\emptyset) = 0$.

The monotonic cover game for an SN-problem is defined as follows:

Definition 2. [3,5] Let $\hat{c}$ be the mst-game for an SN-problem $(V \cup \{*\}, E, w, N, a)$. A function $c : 2^N \to \mathbb{R}_+$ is said to be the monotonic cover of a game $\hat{c}$ or the monotone cover game (mc-game) for the SN-problem $(V \cup \{*\}, E, w, N, a)$ if it satisfies

$$c(S) = \min_{S \subseteq T \subseteq N} \hat{c}(T), \qquad \forall S \subset N.$$

From the definition of monotonic cover, $c(S) \leq \hat{c}(S)$ holds for any $S \subset N$. Hence, it is apparent that $Core(c) \subseteq Core(\hat{c})$ holds.

3 Conventional Solution Concepts

The Bird tree allocation was defined as a solution concept for games in constructing networks.

Definition 3. [1,3,5] Let Γ be a minimum spanning tree w.r.t. an SN-problem $(V \cup \{*\}, E, w, N, a)$ satisfying $P = V \backslash A(N) = \emptyset$. A vector $\boldsymbol{l} = (l_1, l_2, \ldots, l_n)$ is said to be the Bird tree allocation for the minimum spanning tree Γ or a Bird tree allocation for the SN-problem if $l_i = w_{p(a(i))a(i)}$ for any $i \in N$.

The following theorem holds.

Theorem 1. *[3,5] If $\boldsymbol{l}$ is a Bird tree allocation for $(V \cup \{*\}, E, w, N, a)$ satisfying $P = V \backslash A(N) = \emptyset$, then $\boldsymbol{l} \in Core(c)$.*

Let $\boldsymbol{l}$ be the Bird tree allocation for Γ w.r.t. an SN-problem $(V \cup \{*\}, E, w, N, a)$ satisfying $P = V \backslash A(N) = \emptyset$. Remove edges one of whose endpoints are $a(i)$ from Γ. By using remaining vertices, construct a minimum spanning tree for $V \backslash \{a(i)\}$. Let c_r^{-i} denote the cost of the first new edge from r to the source in the constructed minimum spanning tree for $V \backslash \{a(i)\}$. Preparatory to the definition of demand operation by i, define β_r for $r \in F(i)$ by

$$\beta_r = \begin{cases} c_r^{-i}, & \text{if } \sum_{k \in F(i)} c_r^{-i} \leq l_i + \sum_{k \in F(i)} l_k, \\ \alpha_r l_i + l_r, & \text{otherwise,} \end{cases}$$

where $0 \leq \alpha_r \leq (c_r^{-i} - l_r)/l_i$ and $\sum_{r \in F(i)} \alpha_r = 1$. The cost β_r, $r \in F(i)$, represents the cost which is needed for $r \in F(i)$ if player i does not cooperate.

Players directly connected to the supplier does not receive benefit by forming the grand coalition in the Bird tree allocation, although his/her cooperation is important for constructing a link between the supplier and other players in many cases. In other words, although predecessors should have an advantage over followers in the cost allocation, the Bird tree allocation does not give predecessors an advantage. To improve this point, assume that each player will demand that their adjacent followers should share more cost for him/her than that in the Bird tree allocation. Then we had demand operations as follows:

Definition 4. [9]Let $i \in N$ and let $\boldsymbol{y}$ be an allocation. For $r \in N$ and for a minimum spanning tree Γ w.r.t. an SN-problem without public vertices, let

$$\delta_r^i(\boldsymbol{y}) = \begin{cases} \beta_r, & \text{if } r \in F(i), \\ y_r - \sum_{k \in F(i)} (\delta_k^i(\boldsymbol{y}) - y_k), & \text{if } r = i, \\ y_r, & \text{otherwise.} \end{cases}$$

The operation $\boldsymbol{\delta}^i$ which associates the vector $\boldsymbol{\delta}^i(\boldsymbol{y}) = (\delta_1^i(\boldsymbol{y}), \delta_2^i(\boldsymbol{y}), \ldots, \delta_n^i(\boldsymbol{y}))$ with each allocation $\boldsymbol{y}$ is said to be the demand operation by player i in Γ.

Theorem 2. *[9]Let $(V \cup \{*\}, E, w, N, a)$ be an SN-problem satisfying $P = V \backslash A(N) = \emptyset$ and let $\boldsymbol{l}$ be the Bird tree allocation for Γ. If $\boldsymbol{\delta}^i$ is the demand operation by a player i associated with Γ, then $\boldsymbol{\delta}^i(\boldsymbol{l}) \in Core(c)$ for any $i \in N$.*

By applying demand operations by players in a coalition recursively, we obtained a demand operation by the coalition[9]. We showed that any obtained allocation through a operation by a coalition also belongs to the core of the corresponding mc-game.

Granot et al. [4] investigated properties of games in constructing networks with public vertices. They showed the following condition for their core to be nonempty.

Theorem 3. *[4] There exists a minimum spanning tree w.r.t.* $(V \cup \{*\}, E, w, N, a)$ *that spans all vertices and that no two public vertices are adjacent therein. The core of the associated mc-game is not empty.*

4 A New Concept of Demand Operations

In this section, we define a new concept of demand operation. First, consider the case $P = \emptyset$. Let $\boldsymbol{l}$ be the Bird tree allocation for Γ w.r.t. a given SN-problem. As the case for the definition of demand operation by i, remove all edges ones of whose endpoints are $a(i)$ for $i \in S$. By using remaining vertices, construct a minimum spanning tree for $V \backslash A(S)$. Let c_r^{-S} denote the cost of the first edge from r to the source in this minimum spanning tree for $V \backslash A(S)$. We define β_r w.r.t. Γ as follows:

$$\beta_r = \begin{cases} c_r^{-S}, & \text{if } \sum_{k \in F(S)} c_k^{-S} \leq \sum_{i \in S} l_i + \sum_{k \in F(S)} l_k, \\ \alpha_r l_i + l_r, & \text{otherwise,} \end{cases}$$

where $0 \leq \alpha_r \leq (c_r^{-S} - l_r)/l_i$ for any $r \in F(S)$ and $\sum_{r \in F(S)} \alpha_r = 1$. Here, we define a allocation by using a concept of demand operation.

Definition 5. The vector $\boldsymbol{\delta}^S = (\delta_i^S)_{i \in N}$ is said to be a demand allocation by S for a minimum spanning tree Γ w.r.t. an SN-problem if it satisfies

$$\delta_i^S = \begin{cases} \beta_i, & \text{if } i \in F(S), \\ l_i - \dfrac{l_i}{\sum_{j \in S} l_j} \sum_{r \in F(S)} (\delta_r^S(\boldsymbol{l}) - l_r), & \text{if } i \in S, \\ l_i, & \text{otherwise.} \end{cases}$$

We obtain a sufficient condition for the demand allocation to belong to the core of the corresponding mc-game.

Theorem 4. *Let* $(V \cup \{*\}, E, w, N, a)$ *be an SN-problem satisfying* $P = V \backslash A(N) = \emptyset$. *A demand allocation by S for a minimum spanning tree* Γ *w.r.t. an SN-problem is included in the core of the corresponding mc-game if the following holds:*

$$\hat{c}(T) \geq \hat{c}(T \backslash S) + \max\left\{ \sum_{k \in S \cup F(S)} l_k - \sum_{l \in F(S)} c_l^{-S}, 0 \right\} \frac{\sum_{m \in S \cap T} l_m}{\sum_{m' \in S} l_{m'}}, \qquad \forall T \subseteq N.$$

By using Theorem 4, we shall show a new condition for the core of the mc-game associated with an SN-problem with public vertices to be nonempty as follows:

Theorem 5. *For a given $(V \cup \{*\}, E, w, N, a)$, assume that there exists a coalition $S \supseteq P$ satisfying the following conditions.*

1. $\hat{c}(T) \geq \hat{c}(T \backslash S)$ *for any* $T \subseteq N$,
2. *There exists a spanning tree for the given SN-problem which spans* $V \cup \{*\}$, *and for the corresponding Bird tree allocation and the coalition* S, $\sum_{i \in S \cup F(S)} l_i \leq \sum_{i \in F(S)} c_i^{-S}$ *holds.*

Then the core of the corresponding mc-game is not empty. In fact, the demand allocation by S *for the minimum spanning tree* Γ *is an element of the core of the corresponding mc-game, i.e.,* $\boldsymbol{\delta}^S \in Core(c)$.

5 Conclusion

We have discussed coalitional games in constructing networks. The demand allocation, which can be considered as a new version of demand operation by a coalition, has been given. A sufficient condition that demand allocations belong to the cores of coalitional games in constructing networks has been shown. By using this theorem, we have obtained a sufficient condition for the cores of coalitional games in constructing networks with public vertices to be nonempty.

References

1. Bird, C. G., (1976) On cost allocation for a spanning tree: a game theoretic approach, Networks, **6**, 335-350
2. Claus, A., Kleitman, D. J. (1973) Cost allocation for a spanning tree, Networks, **3**, 289-304
3. Curiel, I. (1997) Cooperative Game Theory and Applications: Cooperative Games Arising from Combinatorial Optimization Problems, Kluwer Academic Publishers, Netherlands
4. Granot, D., Maschler, M. (1998) Spanning network games, Int. J. of Game Teory, **27**, 467–500
5. Granot, D., Huberman, G. (1981) Minimum cost spanning tree games, Mathematical Programming, **21**, 1-18
6. Granot, D., Huberman, G. (1984) On the core and nucleolus of minimum cost spanning tree games, Mathematical Programming, **29**, 323-347
7. Kuipers, J., (1993) On the core of information graph games, Int. J. of Game Theory, **21**, 339–350
8. Kuipers, J. (1993) Minimum cost forest games, Int. J. of Game Theory, **26**, 367–377
9. Tsurumi, M., Minamiura, T., Tanino, T., Inuiguchi, M. (2000) Demand Operations in Minimum Spanning Tree Games, Game Theory and Applications, **5**, 142–155

Multi-Objective Facility Location Problems in Competitive Environments

Takeshi Uno[1], Hiroaki Ishii[2], Seiji Saito[2], and Shigehiro Osumi[3]

[1] Department of Applied Physics, Graduate School of Engineering, Osaka University, 2-1, Yamadaoka, Suita, Osaka, 565-0871 Japan
[2] Department of Information and Physical Sciences, Graduate School of Information Science and Technology, Osaka University, 2-1, Yamadaoka, Suita, Osaka, 565-0871 Japan
[3] Faculty of Design, Kobe Design University, 8-1-1, Gakuennishi-machi, Nishi-ku, Kobe, 651-2196 Japan

Abstract. In this paper we propose a new model of competitive facility location on a plain. We consider the two objectives that are maximization of reward for a firm and convenience of facility location for customers. We formulate the model as a single facility location problem under the condition that a competitive facility has been already located. We construct efficient algorithms to solve the problem, and give numerical experiments to illustrate our algorithms.

1 Introduction

The medianoid problem, which was suggested by Hakimi [2], deals with the location of new facilities under the condition that competitive facilities have been already located [5]. Hakimi [2] considered the problem on a network that all demand points are nodes in. On the other hand, Drezner [1] considered the problems on the plane which includes demand points.

In these problems, it is usually assumed that each customer on demand points selects the closest facility to him/her. Karkazis [4] extended Hakimi's medianoid problems by assuming that customers decide their using facilities based on the following two generalized criteria: a "distance criterion" regarding the vicinity of the facility and a "level criterion" regarding the type of the facility. Uno et al. [6] extended Drezner's medianoid problems by assuming attractive function of the distance and the level.

In most competitive facility location models, main objective of each decision maker is maximization of its reward obtained by customers. However, the location of facilities only based on such an objective is often inconvenient for customers. The models with multi-objective for firms and customers on a network have been studied [7,8].

In this paper we propose a new model of competitive facility location with considering convenience of the location for customers. We extend Uno et al.'s medianoid problems [6] to multi-objective problems for a firm and customers. We construct algorithms to solve the problems and give numerical experiments to illustrate the algorithms.

The remainder of this paper is organized as follows. In Section 2, we introduce the medianoid problem of Uno et al. [6] and show some features of solutions for the problem. In Section 3, we formulate a multi-objective problem for a firm and customers that extends the problem in Section 2. In Section 4, we propose efficient algorithms for the problem in Section 3. In Section 5, we present results of computational experiments for an example of our models. Finally, in Section 6, we make mention of conclusions.

2 Medianoid problem with single objective

There are n demand points on the plane $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$, which are indicated in $I = \{1, \cdots, n\}$. Let $v_i = (x_i, y_i) \in \mathbf{R}^2$ denote a site of demand points $i \in I$, and $w_i \in (0, \infty)$ a sum of purchasing powers with customers on demand points $i \in I$.

It is assumed that a competitive facility, denoted by A, has already located on $\mathbf{R}^2$. Let $u_A = (x_A, y_A) \in \mathbf{R}^2$ and $l_A \in \{1, \cdots, L\}$ denote a site and a quality level of A, respectively, where L is a natural number that means maximal quality level of facility.

Now we consider the location of a new facility, denoted by B. Let $u_B = (x_B, y_B) \in \mathbf{R}^2$ and $l_B \in \{1, \cdots, L\}$ denote a site and a level of B, respectively. Let $T \equiv \mathbf{R}^2 \times \{1, \cdots, L\}$ denote a set of solutions for B.

We consider the following two functions about levels. Let $k : \{1, \cdots, L\} \to [1, \infty)$ denote a function which estimates attractiveness of facilities such that $\infty > k(1) > \cdots > k(L) \geq 1$. Let $C : \{1, \cdots, L\} \to [0, \infty)$ denote a function which estimates a building cost of a facility such that $0 \leq C(1) < \cdots < C(L) < \infty$.

Let $d(p^1, p^2)$ denote a distance between two points p^1, p^2 of $\mathbf{R}^2$. In this problem, for demand point $i \in I$, customers at v_i always use A if

$$k(l_A) d(u_A, v_i) \leq k(l_B) d(u_B, v_i), \tag{1}$$

otherwise, the customers always use B. Let N_B denote a set of indices of demand points whose purchasing powers are captured by B.

In this problem, the objective of decision maker is to maximize his/her reward obtained from B. It is assumed that the reward is linear for captured purchasing powers. Let $\alpha \geq 0$ denote a coefficient about the reward for B. The reward obtained by facility B is represented as

$$r(u_B, l_B) = \alpha \cdot \sum_{i \in N_B} w_i - C(l_B). \tag{2}$$

Therefore the medianoid problem is formulated as follows:

$$\begin{aligned} P_R : \ \max \ & r(u_B, l_B) && (3) \\ \text{s.t.} \ & (u_B, l_B) \in T, && (4) \\ & N_B = \{ i \in I \mid k(l_A) d(v_i, u_A) > k(l_B) d(v_i, u_B) \}. && (5) \end{aligned}$$

In order to solve Problem (P_R), we need to examine B's site to all levels of B. For levels of A and B, the following three cases are considered.

In cases that $l_B = l_A$, the problem to find optimal sites for B is equivalent to the medianoid problem of Drezner [1]. Then the following theorem is given as Theorem 1 in Drezner [1]:

Theorem 1. *Let $l_B = l_A$. One of the optimal location for a follower firm is infinitesimally close to u_A but not on u_A.*

In cases that $l_B > l_A$, the following theorem is given as Proposition 2 in Uno et al. [6]:

Theorem 2. *Let $l_B > l_A$. Then one of the optimal sites for B is common to A's site.*

In cases that $l_B < l_A$, algorithm to find one of optimal sites for B is given as Algorithm 2 in Uno et al. [6]. The complexity of the algorithm is $O(|I|^3)$. Note that optimal sites for B usually separate from u_A.

In the former two cases location of A and B is usually inconvenient for customers. In the following section, we extend Problem (P_R) to a multi-objective problem for a firm and customers.

3 Medianoid problem with multi-objective

First we formulate a single objective problem for customers. From the point of view of customers on demand point v_i, we represent degree of inconvenience of facility $F \in \{A, B\}$ as $k(l_F)d(v_i, u_F)$. This means that a facility is convenient for customers if its level is high and its site is near to the customers. We assume that customers only use a facility whose degree of inconvenience is minimal in A or B. Then we represent degree of improvement in convenience for all customers on demand points v_i after B is located as

$$f_i(u_B, l_B) \equiv \begin{cases} w_i \cdot (k(l_A)d(v_i, u_A) - k(l_B)d(v_i, u_B)), & \text{if } i \in N_B, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Therefore we formulate location problem for customers as the following maximizing problem:

$$P_C: \ \max \sum_{i \in I} f_i(u_B, l_B) \tag{7}$$

$$\text{s.t.} \ (u_B, l_B) \in T. \tag{8}$$

Next we combine Problem (P_C) with Problem (P_R) in order to formulate multi-objective problem for a firm and customers. We assume that the B can be shifted in $\mathbf{R}^2$ under the condition that the decision maker can obtain maximal reward. We denote a feasible set of Problem (P_R) as

$$T_B \equiv \arg \max_{(u_B, l_B) \in T} \{r(u_B, l_B)\}. \tag{9}$$

Therefore we formulate multi-objective problem for a firm and customers as follows:

$$P: \max \sum_{i \in I} f_i(u_B, l_B) \tag{10}$$

$$\text{s.t. } (u_B, l_B) \in T_B, \tag{11}$$

4 Algorithm for competitive facility location problems

First we divide the feasible set of Problem (P) according to sets of customers which use B. Let S_B denote a total set of N_B if $(u_B, l_B) \in T_B$. Let $\bar{N}_B \in S_B$ be given and $T_B(\bar{N}_B)$ denote a feasible set of Problem (P) such that a set of B's customers is $\bar{N}_B$. Since sum of terms about A in (6) for all demand points in $\bar{N}_B$ is fixed, we can estimate improvement in convenient by using only a term about B in (6). Then in a part of feasible set $\bar{N}_B$, the following problem is equivalent to Problem (P):

$$P_{\bar{N}_B}: \min \sum_{i \in \bar{N}_B} w_i \cdot k(l_B) d(u_B, v_i) \tag{12}$$

$$\text{s.t. } (u_B, l_B) \in T_B(\bar{N}_B), \tag{13}$$

$$\bar{N}_B \text{ is given.} \tag{14}$$

We solve Problem $(P_{\bar{N}_B})$ for all elements in S_B so that we can find optimal solution for Problem (P). Problem $(P_{\bar{N}_B})$ is an application of Weber problems [9], which are one of classical optimal location problems whose objective is to find the facility location which is convenient for customers.

Now we construct an efficient algorithm to solve Problem $(P_{\bar{N}_B})$. For $\bar{N}_B$, a feasible set of Problem $(P_{\bar{N}_B})$ is represented as

$$T_B(\bar{N}_B) = \bigcap_{i \in \bar{N}_B} \{(u_B, l_B) \in T_B \mid k(l_B) d(u_B, v_i) < k(l_A) d(u_A, v_i)\}. \tag{15}$$

From the second term in (2), the optimal level for Problem $(P_{\bar{N}_B})$ is given uniquely. From (12), objective function of Problem $(P_{\bar{N}_B})$ is convex. From (15), a set of site in $T_B(\bar{N}_B)$ is an intersection of convex sets. Therefore Problem $(P_{\bar{N}_B})$ is a convex programming problem. In Section 2, we have already introduced the method of finding feasible interior-point solutions for Problem $(P_{\bar{N}_B})$ by dividing into the three cases. We use quasi-Newton method [3] to solve Problem $(P_{\bar{N}_B})$ for each $\bar{N}_B \in S_B$.

5 Numerical experiments

We illustrate our medianoid problems by the following numerical example. Distribution of customers is given in Table 1. For a level of facility that the

Table 1. Distribution of customers

Index i	Site $u_i = (x_i, y_i)$	Purchasing power w_i
1	(0.00, 2.00)	300
2	(2.00, 5.00)	300
3	(4.00, 0.00)	200
4	(5.00, 3.00)	400
5	(8.00, 5.00)	100
6	(9.00, 2.00)	200

firm can locate, $L = 3$. A has already been located on v_4 and its level is that $l_A = 2$. Functions that describe attractiveness and cost of building are given by Table 2.

Table 2. Numerical data of $k(l_F)$ and $C(l_B)$

Level of facilities $l_F, F \in \{A, B\}$	Attractive function $k(l_F)$	Cost function $C(l_B)$
1	4.00	100
2	2.00	200
3	1.00	400

First we solve Problem (P_R) introduced in Section 2. If $0 \leq \alpha \leq 3/2$, optimal level of B is that $l_B = 3$, and from Theorem 2, one of optimal sites is on v_4. Then B can capture purchasing powers from all demand points except demand point No.4, and B's reward is $1100 - \alpha \cdot 400$. If $3/2 \leq \alpha \leq 2$, optimal level of B is that $l_B = 2$, from Theorem 1, one of optimal sites is infinitesimally close from the left side of u_A to u_A but not on u_A. Then B can capture from demand points No.1, 2, and 3, and B's reward is that $800 - \alpha \cdot 200$. If $\alpha \geq 2$, optimal level of B is that $l_B = 1$, and we use Algorithm 2 in Uno et al. [6] to find one of optimal sites that $u_B = (1.17, 3.76)$ in a line segment between v_1 and v_2. Then B can capture from demand points No.1 and 2, B's reward is $600 - \alpha \cdot 100$.

Secondly, in order to solve Problem (P), we use our algorithm proposed in Section 4. If $0 \leq \alpha \leq 3/2$, an optimal site for Problem (P) is that $u_B = (2.65, 3.01)$. Then improvement in convenience is 4.86×10^3. If $3/2 \leq \alpha \leq 2$, an optimal site for Problem (P) is that $u_B = (2.73, 1.47)$. Then improvement in convenience is 2.62×10^3. If $\alpha \geq 2$, optimal sites for Problem (P) are in the line segment between v_1 and v_2. This means that the optimal site for Problem (P_R) is one of optimal sites for Problem (P). Then improvement in convenience is 8.96×10^2.

From the result of our numerical experiments, we know that the less the weight of building cost to captured purchasing powers is, the more the facility location is convenience for customers. However, in cases that level of B is higher than that of A, from Theorem 1 and 2, we know that the location of facilities only based on reward is not so convenient for customers. These mean importance of estimation of convenience for customers in competitive facility location problems.

6 Conclusions

In this paper we have extended Uno et al.'s medianoid problems [6] to multi-objective problems for a firm and customers. We have constructed efficient algorithms for solving the problems, and as the result of numerical experiments for an example of our model, we have obtained facility locations that are convenient for customers.

We have considered medianoid problems but not centroid problems, which was also suggested by Hakimi [2]. The problems deal with location of new facilities under the condition that competitive facilities will be located after [5]. Studies of centroid problems with multi-objective are future researches.

References

1. Drezner, Z. (1982) Competitive Location Strategies for Two Facilities. Regional Science and Urban Economics **12**, 485–493
2. Hakimi, S. L. (1983) On Locating New Facilities in a Competitive Environment. European Journal of Operational Research **12**, 29–35
3. Ibaraki, T., Fukushima, M. (1993) Methods of optimization, Kyoritsu Shuppan Co., Ltd. (in Japanese)
4. Karkazis, J. (1989) Facilities Location in a Competitive Environment: A Promethee Based Multiple Criteria Analysis. European Journal of Operational Research **42**, 294–304
5. Osumi, S. (1997) Modelling and Analysis of Competitive Facility Location Problem. Doctor's thesis of Osaka University
6. Uno, T., Ishii, H., Saito, S., Osumi, S. A location model in a competitive environment considering quality levels of facilities. Computers and Operations Research (in submission)
7. Uno, T. (1999) Facility location problems with multi-objective. Master's thesis of Osaka University (in Japanese)
8. Uno, T., Ishii, H., Saito, S., Osumi, S. (2000) A location model for obnoxious facilities in a competitive environment. Kokyuroku of Research Institute for Mathematical Sciences, Kyoto University: Mathematical Science of Optimization **1174**, 71–82 (in Japanese)
9. Weber, A. (1909) Über Den Standort Der Industrien, 1. Teil: Reine Theorie Des Standortes, Tübingen, Germany. English Translation: on the Location of Industries, University of Chicago Press, Chicago, IL, 1929. (English Translation by C. J. Friedeich (1957), Theory of the Location Of Industries, Chicago University Press, Chicago.)

Solving Portfolio Problems Based on Meta-Controled Boltzmann Machine*

Junzo Watada[1]** and Teruyuki Watanabe[1]

Osaka Institute of Technology, 5-16-1 Omiya Asahi, Osaka 535-8585 Japan

Abstract. It is important that the limited amount of investing funds should be efficiently allocated to many stocks so as to reduce its risk. This problem is formulated as a mixed integer programming problem. However, it is not so easy to solve the mixed integer programming problem because of its combinatorial nature. Therefore, an efficient approximate method is required to solve a large–scale mixed integer programming problem. In this paper we propose a Meta-controlled Boltzmann machine to obtain an approximate solution of the large–scale mixed integer programming problem.

1 Introduction

A mean-variance approach to portfolio selection problem has been originally proposed by H. Markowitz [3]. It, based on time-series data of return rate, theoretically decides the best investing rate to each of stocks which minimizes the risk or the variance of the profits in keeping the least expected return rate that a decision maker expects. The objective of the Markowitz' model is to reduce its risk in allocating the amount of investing funds to many stocks.

In this paper, we propose a Meta-controlled Boltzmann machine [7], based on a Two-layered Boltzmann machine which is proposed by J. Watada et. al [4], to solve the portfolio selection problem which limits the number of invested stocks. The meta-controlled Boltzmann machine consists of a Hopfield network [2] as the Meta-controlling layer and a Boltzmann machine [1] as the lower layer. The Meta-controlling layer supervises the subordinated lower layer to obtain the best portfolio within the optimal combination of invested stocks and the lower layer decides the optimal investment rate over the limited number of stocks supervised by the Meta-controlling layer. This model deletes the units of the lower layer which are not selected in the Meta-controlling layer in the execution. Then the lower layer is restructured by using the selected units. Executing the meta-controlled Boltzmann machine according to the above-mentioned algorithm, the meta-controlled Boltzmann machine converges more efficiently than a conventional Boltzmann machine.

* This research was supported in part by Grant-in Aid for Scientific Research(C-2); Grant No.11680459 of Ministry of Education of Science, Sports and Culture.

** All enquiries can be sent to Professor Junzo Watada/ School of Industrial Management, Osaka Institute of Technology, 5-16-1 Omiya Asahi, Osaka 535-8585 Japan / E-mail: junzow@osb.att.ne.jp

In this paper, we evaluate the efficiency of the Meta-controlled Boltzmann machine employing various size of real data.

2 Portfolio Selection Problem

The mean-variance approach to the portfolio selection problem was originally proposed by H. Markowitz [3]. In the formulation of portfolio selection model, H. Markowitz started his discussion with the assumption that almost all decision makers have aversion to risk even if its return may be obtained less in decreasing the risk. Since the risk is estimated under the condition of fixing the expected return rate, the decision maker cannot be fully satisfied with its solution.

In a real problem, it is important to find the optimal selection of invested stocks out of many feasible stocks in a market, because of the limit amount of funds to invest into a stock market. This problem is formulated by a mixed integer programming problem as follows:
FORMULATION 1.

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{n} \mu_i m_i x_i \\ \text{minimize} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij} m_i x_i m_j x_j \\ \text{subject to} \quad & \sum_{i=1}^{n} m_i x_i = 1, \quad \sum_{i=1}^{n} m_i = S \\ & m_i \in \{0,1\}, \quad x_i \geq 0 \quad (i = 1, 2, \ldots, n) \end{aligned}$$

where σ_{ij} denotes a covariance between stocks i and j, μ_i an expected return rate of stock i and x_i an investing rate to stock i, respectively. m_i denotes a selection variable out of investing stocks which takes one for a selected stock and zero for a not-selected stock and S a total number of selected stocks.

However, it is not so easy to solve the mixed integer programming problem because of its combinatorial nature. Therefore, it needs an efficient approximate method to solve a large–scale mixed integer programming problem.

3 Energy Functions for Meta-controlled Boltzmann Machine

The energy function, which is proposed by J. Hopfield, is written in the following equation:

$$E = -\frac{1}{2} \sum_{ij=1}^{n} w_{ij} V_i V_j + \sum_{i=1}^{n} \theta_i V_i, \tag{1}$$

where w_{ij} is a weight between neuron i and j, θ_i is a threshold of neuron i and V_i is output value of unit i, respectively. J. Hopfield has shown that this energy function monotonously decreases as the neural network is executed. There is a possibility that this energy function converges to one of local minima. In the case of a Boltzmann machine, the energy function can increase with a minute probability. Therefore, the energy function hardly falls into a local minimum. Therefore, transforming the objective function of a portfolio selection problem into an energy function of the Boltzmann machine, the Boltzmann machine can solve the portfolio selection problem as its approximate solution. However the computing time of a conventional Boltzmann machine becomes so much large as the number of units increases. Therefore, it is necessary to delete the useless units in order to shorten the computing time.

As the investable fund is limited in real stock investment, the fund should be allocated over the small number of stocks out of a huge stock market. Therefore, we propose a Meta-controlled Boltzmann machine to solve the portfolio selection problem with a huge number of stocks, which is illustrated in Fig. 1.

We can formulate the energy functions of the Meta-controlled Boltzmann machine as in equations (2) and (3). Each unit of the Boltzmann machine corresponds to each stock for investment, and the number of units and the number of stocks should be equal. The investing rate to each stock is obtained as the value which divides the output value of each unit by the total of output value. The covariance between stocks i and j, σ_{ij}, corresponds to weight between neurons i and j, w_{ij}. The expected return rate of stock i, μ_i, corresponds the threshold of neuron i, θ_i. On the condition that the energy function should be minimized, the energy functions are written as follows:

Meta-Controlling Layer

$$E_u = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}s_is_j + K_u\sum_{i=1}^{n}\mu_is_i, \tag{2}$$

Lower Layer

$$E_l = -\frac{1}{2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_ix_j + \sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j\right\} + \sum_{i=1}^{n}x_i + K_l\sum_{i=1}^{n}\mu_ix_i, \tag{3}$$

where K_u and K_l are a weight of the expected return rate for each layer and s_i is an output value of unit i of the Meta-controlling layer.

The output value of the Meta-controlling layer, s_i, is 0 or 1. If stock i is selected, the output value of unit i is 1. The output value of the lower layer is investing ratio to each stock. In the Meta-controlling layer if K_u is set to a larger value, the selected number of stocks will increase. In the lower layer if K_l is set to a smaller value, we will obtain the solution nearer to

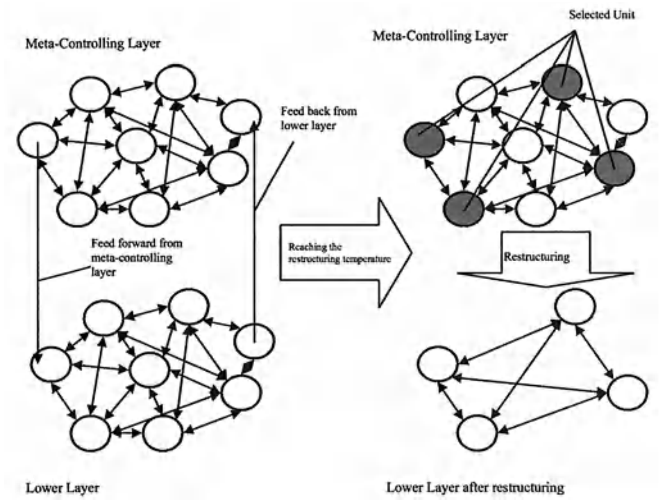


Fig. 1. Meta-controlled Boltzmann Machine

Step 1.	Set each parameter to an initial value.
Step 2.	Input the value of $K_u and K_l$.
Step 3.	If output value of the Meta-controlling layer is 1, increase the value of the corresponding lower layer's unit.
Step 4.	If the output value is sufficiently large, increase the value of the corresponding Meta-controlling layer's unit.
Step 5.	Iterate from Step 3 to 4 until the temperature reaches the restructuring temperature.
Step 6.	Restructure the lower layer using the selected units.
Step 7.	Execute the lower layer until reaching at the termination.

Fig. 2. Algorithm of a Meta-controlled Boltzmann machine

the minimum risk solution. Therefore, as we change each parameter K, we will obtain various investing pattern according to the aspiration of a decision maker's in the same as a fuzzy portfolio selection model which is proposed by J. Watada et al [5,6]. Algorithm of the meta-controlled Boltzmann machine is shown in Fig. 2. Restructure the lower layer can make its computing time much shorter to reach at the termination.

4 Numerical Example

In this section, we show the effectiveness of a Meta-controlled Boltzamnn machine employing various size of real data from 10 stocks to 1286 stocks in Tokyo Market. The results are shown in Table 1 and Fig. 3. Computing efficiency is given by the following equation:

$$C_e = \frac{t_{Meta}}{t_{conv.}} \times 100 \tag{4}$$

Table 1. Comparison of Meta-controlled Boltzmann Machine with conventional Boltzmann Machine

Number of units	CPU Time (sec.)		Computing Efficiency (%)	Expected return rate		Risk		Selected units	
	Conventional	Meta-controlled		Conventional	Meta-controlled	Conventional	Meta-controlled	Conventional	Meta-controlled
10	7.28	6.20	85.2	0.00138	0.00138	0.00088	0.00088	6	3
20	9.03	6.44	71.3	0.00090	0.00091	0.00073	0.00073	4	3
40	13.08	7.38	56.4	0.00081	0.00080	0.00053	0.00052	5	5
80	21.81	8.56	39.2	0.00137	0.00138	0.00063	0.00063	8	7
160	39.50	9.96	25.2	0.00306	0.00307	0.00057	0.00058	4	6
320	103.56	19.18	18.5	0.00402	0.00400	0.00044	0.00042	5	9
640	229.43	38.02	16.6	0.00738	0.00739	0.00095	0.00096	6	15
1286	491.49	87.96	17.9	0.00736	0.00742	0.00092	0.00099	6	15

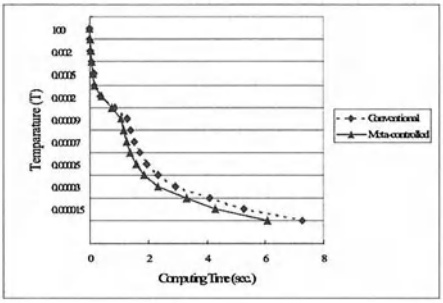
where C_e denotes computing efficiency, t_{Meta} a computing time of the Meta-controlled Boltzmann machine and $t_{conv.}$ a computing time of a conventional Boltzmann machine, respectively. In Table 1, it is shown that the computing time of the Meta-controlled Boltzmann machine is drastically shorter than a conventional Boltzmann machine. The reason of this result is because the Meta-controlled Boltzmann machine deletes useless units in step of restructuring in the execution. On the other hand in the case of a conventional Boltzmann machine, it is computed employing all units until reaching the termination. Therefore, the computing time of the Meta-controlled Boltzmann machine is effectively and efficiently shorter than a conventional Boltzmann machine. Comparing with the value of computing efficiency, it shows that the Meta-controlled Boltzmann machine is more efficient, so that an initial unit value is large. Comparing between expected return rate and risk, we can obtain satisfactorily a good result employing the Meta-controlled Boltzmann machine.

5 Concluding Remarks

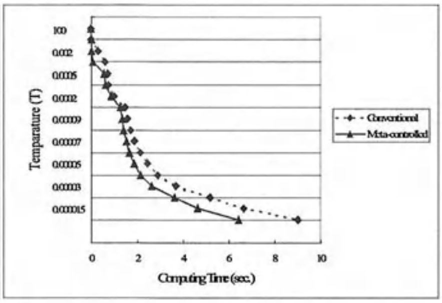
In this paper, we proposed the Meta-controlled Boltzmann machine, based on a Two-layered Boltzmann machine which is proposed by J. Watada et. al [4], to solve the portfolio selection problem which limits the number of invested stocks. The result of the numerical example stresses that the solution with a limited number of selected stocks can be obtained more efficiently employing the Meta-controlled Boltzmann machine.

References

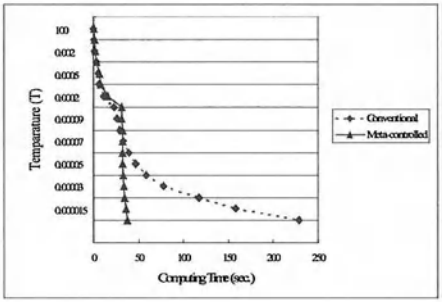
1. Ackley, D. H., Hinton, G. E., Sejnowski, T. J. (1985) "A Learning Algorithm for Boltzmann machines," *Cognitive Science*, **9**, 147–169.
2. Hopfield, J. J. (1982) "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. National Science*, 2554–2558
3. Markowitz, H. (1952) "Portfolio Selection," *Journal of Finance*, **7**, 77-91



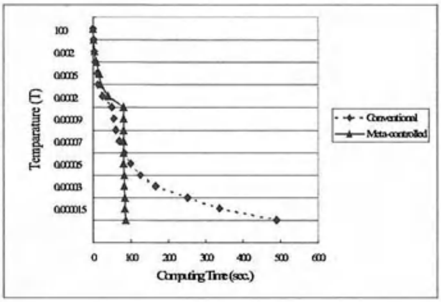
a) Number of units 10



b) Number of units 20



c) Number of units 640



d) Number of units 1286

Fig. 3. Comparison of the computing time between a conventional Boltzmann Machine and a Meta-controlled Boltzmann Machine

4. Watada, J., Oda, K. (1999) "Formulation of a Two-layered Boltzmann Machine for Portfolio Selection," *International Journal of Fuzzy Systems*, **2**, **1**, 39-44
5. Watada, J. (2001) Fuzzy Portfolio Model for Decision Making in Investment, *Dynamical Aspects in Fuzzy Decision Making*, ed. K. Yoshida, Physica-Verlag, Springer, 141-162
6. Watada, J. (2002) "On Recent Development of Soft-Computing Approaches to Portfolio Selection Problems," *Proceedings, 2nd European Conference on Intelligent Systems and Technologies (ECIT2002)*
7. Watanabe, T., Watada, J. (to appear) "A Meta-Controlled Boltzmann Machine for Rebalancing Portfolio Selection," *Central European Journal of Operations Research*

Tradeoff directions and dominance sets

Petra Weidner

FB PMF, FH Hildesheim/Holzminden/Göttingen, Von-Ossietzky-Str. 99,
D-37085 Göttingen, Germany
petra.weidner@pmf.fh-goettingen.de

1 Introduction

In this paper, we provide a short survey of tradeoff concepts and present a scalarization the parameters of which allow one to calculate the tradeoffs as efficient points of the contingent cone in its optimal solution, if certain differentiability assumptions are fulfilled at this efficient point. The scalarization is based on the idea to look for vectors being efficient w.r.t. some widened dominance cone.

From now on, we will assume $F = f(X)$ to be the feasible point set in the objective space $\Re^p$ of the multicriteria optimization problem

$$\begin{array}{lcl} f_1(x) & \to & \min \\ & \vdots & \\ f_p(x) & \to & \min \\ & x \in X, & \end{array}$$

where X stands for some arbitrary set of feasible alternatives and $f := (f_1, \ldots, f_p)^T$ consists of the real-valued functions $f_i : X \to \Re$, $i=1,2,\ldots,p$, $p \geq 2$.

We define the optima of this problem as follows:

Definition 1.

1. $y \in F$ is said to be an *efficient element* of F iff $F \cap (y - \Re^p_+) = \{y\}$. Eff(F) will denote the set of all efficient elements of F.
2. $y \in F$ is a *weakly efficient element* of F iff $F \cap (y - \text{int}\, \Re^p_+) = \emptyset$.
3. $y^o \in F$ is called a *properly efficient element* of F (in the sense of Geoffrion) iff $\exists K > 0 : \forall y \in F \text{ with } y_i < y_i^o$ for some $i \in \{1,2,\ldots,p\}$:

$$\exists j \in \{1,2,\ldots,p\} \setminus \{i\} : y_i^o - y_i \leq K(y_j - y_j^o).$$

The set of all properly efficient elements of F is expressed by G-Eff(F).

Moreover, we define optimality with respect to an arbitrary set $D \subseteq \Re^p$ that can be an aggregation of the decision maker's dominance directions or – in the sections to come – a set related to the scalarizing problem.

Definition 2. Let $D \subseteq \Re^p$ be an arbitrary set.
$y \in F$ is said to be an *efficient element* of F w.r.t. D iff $F \cap (y - D) \subseteq \{y\}$.

Eff(F,D) will specify the set of these elements.

Obviously, $\mathrm{Eff}(F,\mathfrak{R}^p_+)=\mathrm{Eff}(F)$.

Throughout, cone(F) will denote the cone generated by the set F, cl F the topological closure of the set F, bd F its topological boundary,

$T(F,y^0) := \{d\in \mathfrak{R}^p \mid \text{for all } (t_j)\subset\mathfrak{R} \text{ with } t_j \downarrow 0 \text{ and all } (y^j)\subseteq F \text{ with } y^j \to y^0,$

$\text{there exists some sequence } (d^j)\subset\mathfrak{R}^p \text{ with } d^j \to d \text{ and } y^j + t_j d^j \in F\}$

the tangent cone of F in y^0 and

$K(F,y^0) := \{d\in \mathfrak{R}^p \mid \text{there exist sequences } (t_j)\subset \mathfrak{R} \text{ with } t_j \downarrow 0 \text{ and } (d^j)\subset\mathfrak{R}^p$

$\text{with } d^j \to d \text{ such that } y^0 + t_j d^j \in F\}$

the contingent cone of F in y^0.

2 Tradeoff Concepts

Traditionally, tradeoffs are thougt to be the price one is willing to pay by a worse value in one objective function for improving the value of another one. This view has been extended by introducing vectors that express a combined impairment and a combined improvement of objective function values. In this way the consideration changes from directions parallel to the axes to sets of potential tradeoff vectors.

The classical local tradeoff between objective functions f_i and f_j expresses how much the decision maker is willing to or has to gain in objective i for a unit increase in objective j, if the other objective function values are fixed.

Assuming u to be some explicit or implicit value function that has to be maximized over f(X) such type of definition is given as the marginal rate of substitution or local tradeoff ratio by $\frac{\frac{\partial u(f(x^0))}{\partial f_i(x)}}{\frac{\partial u(f(x^0))}{\partial f_j(x)}}$. The local tradeoff ratio measured by values of relative changes in the objective function values has also been investigated without using a value function as $\frac{\partial f_i(x^0)}{\partial f_j(x^0)}$.

Not being based on infinitesimal properties, the point-to-point tradeoff $t_{ij}(y^0, y) := \frac{y_i^0 - y_i}{y_j - y_j^0}$ between objective functions f_i and f_j for feasible outcomes y^0 and y expresses how many units of the decrease in attribute i are gained for a unit of increase in attribute j, if the other function values y_k coincide with the corresponding function values y_k^0. Without this assumption, $-t_{ij}(y^0, y)$ is called a total tradeoff by Haimes and Chankong (1979), otherwise a partial tradeoff. A total tradeoff turns out to be useful for the comparison of a small number of efficient solutions.

Kaliszewski (1994) modified this definition in such a way that other objective function values are allowed to alter as long as they do not become worse and that the tradeoff compares y^0 with the entire feasible outcome set. He defined the tradeoff $T_{ij}(y^0) := \sup_{y \in Z_j^{>}(y^0)} \frac{y_i^0 - y_i}{y_j - y_j^0}$,

where $Z_j^{>}(y^0) := \{y \in F \mid y_j > y_j^0,\ y_k \leq y_k^0 \text{ for all } k \neq j\}$. This as well as the total tradeoff extends the tradeoff directions from vectors parallel to the coordinate axes to other vectors of interest.

Haimes and Chankong (1979) investigated tradeoff rates in x^0 w.r.t. a direction d in the decision space as $\lim_{\alpha \to 0} \frac{f_i(x^0 + \alpha d) - f_i(x^0)}{f_j(x^0 + \alpha d) - f_j(x^0)}$ and calculated them by means of the Kuhn-Tucker multipliers in the ε-constraint approach. Similar calculations were applied to the weighted Tchebycheff norm method by Yano and Sakawa (1987).

A generalization of this infinitesimal concept was introduced by Henig and Buchanan 1997. They defined the cone of tradeoff directions for F at y^0 as Eff(cl cone($F - y^0$)) under the assumption, that $F + \mathfrak{R}_+^p$ is convex. An extension of this definition for feasible point sets without convexity properties was introduced by Lee and Nakayama 1997 as well as by Miettinen and Mäkelä 2002. They considered the efficient point set of the contingent cone of F at y^0 as the cone of tradeoff directions of F at y^0. This contingent cone is always closed, but not necessarily convex for a nonconvex set F. It is contained in cl cone($F - y^0$) and coincides with it, if F is starshaped at y^0.

In (Kaliszewski 1994, 2000, Kaliszewski and Michalowski 1997) parameters of scalarizing problems are used to compute bounds for tradeoffs in the properly or weakly efficient solution obtained. These scalarizing problems are just of the type ($P_{a,k,C}$) that we will present in the next section with C being polyhedral cones. This was pointed out in (Weidner 1994) in connection with (Weidner 1990).

3 A Scalarization using Widened Dominance Sets

In (Weidner 1990, 2001) a general scalarization was discussed that extends an approach by Pascoletti and Serafini (1984) by replacing the dominance cone D, here $\Re^p_+$, by some parameter-depending set C containing $D\backslash\{0\}$:

$$(P_{a,k,C}): \quad \begin{array}{l} t \rightarrow \min \\ y \in a - clC + tk \\ y \in F, t \in \Re \end{array}$$

with parameters $a, k \in \Re^p, C \subseteq \Re^p$.

In (Weidner 1993) we investigated the following scalarization:

$$(P): \quad \begin{array}{ll} t \rightarrow \min & \\ -w_i y_i - v_i + t \geq 0 & \forall i = 1,\ldots,p \\ \prod_{i=1}^{p} \left(-w_i y_i - v_i + t\right) \geq b & \\ y \in F,\ t \in \Re & \end{array}$$

with parameters $w \in \text{int}\,\Re^p_+,\ b \in \Re_+ \backslash\{0\},\ v \in \Re^p$.

This problem is equivalent to $(P_{a,k,C})$ with

$$C = \left\{ y \in \Re^p \mid \prod_{i=1}^{p} \left(w_i y_i + \sqrt[p]{b} \right) \geq b,\ w_i y_i + \sqrt[p]{b} \geq 0\ \forall i = 1,\ldots,p \right\},$$

$$a_i = -\frac{v_i + \sqrt[p]{b}}{w_i},\ k_i = \frac{1}{w_i}\quad \forall i = 1,\ldots,p.$$

Geometrically, $(P_{a,k,C})$ and thus (P) can be interpreted as follows:
Stick the set $-clC$ to the point a and shift the set $a - clC$ along the line $\{a + tk \mid t \in \Re\}$ until the smallest parameter $t = t_0$ is reached for which the intersection with F is nonempty, that is for which $F \cap (a - clC + t_0 k) \neq \emptyset$. Then t_0 is the optimal value of the scalarizing problem and $F \cap (a - clC + t_0 k)$ is just the set of optimal solutions y of this problem.

If C and k fulfil the conditions

- $\Re^p = \bigcup_{\alpha \in \Re} (bdC + \alpha k)$,
- $intC = \bigcup_{\alpha \in \Re_+ \backslash\{0\}} (bdC + \alpha k)$,
- $clC + int\,\Re^p_+ \subseteq clC$ and
- $bdC + (\Re^p_+ \backslash\{0\}) \subseteq intC$,

then each solution y^0 of $(P_{a,k,C})$ belongs to Eff(F). The conditions are fulfilled for several classes of sets C containing $\Re^p_+ \setminus \{0\}$ in its interior and vectors $k \in \text{int}\Re^p_+$, especially by the polyhedral cones in the scalarizing problems considered by Kaliszewski and by the set C and the vector k appearing in problem (P). In many of these cases the optimal solutions y^0 of $(P_{a,k,C})$ are efficient elements of F w.r.t. C such that C can be interpreted to be an enlargement of the dominance cone $\Re^p_+$, a widened dominance set.

4 Calculation of Tradeoffs

From now on, consider C, a and k to be as specified for problem (P).

C is a closed, strictly convex set with $C + (\Re^p_+ \setminus \{0\}) \subseteq \text{int}C$ and $0 \in \text{bd}C = \left\{ y \in \Re^p \mid \prod_{i=1}^{p} \left(w_i y_i + \sqrt[p]{b} \right) = b,\ w_i y_i + \sqrt[p]{b} \geq 0\ \forall i = 1,\dots,p \right\}$. Since the surface of C is smooth, each tangent cone of C is a halfspace.

Let t_0 be the optimal value and $y^0 \in F$ be an arbitrary optimal solution of (P). Then y^0 belongs to $a - C + t_0 k$ and $T(a - C + t_0 k, y^0) = \{y \in \Re^p \mid n^T y \leq n^T y^0\}$, where a normal n to the set $a - \text{bd}C + t_0 k$ can be determined by the parameters of (P):

$$n_j = w_j \cdot \prod_{\substack{i=1 \\ i \neq j}}^{p} \left(-w_i y_i^0 - v_i + t_0 \right) \quad \forall j = 1,\dots,p. \tag{1}$$

Moreover, the formulation of (P) immediately yields $n \in \text{int}\Re^p_+$.

If F is closed and $K(F,y^0)$ turns out to contain some hyperplane H, then $H = \{y \in \Re^p \mid n^T y = n^T y^0\}$ and $K(F,y^0) \subseteq \{y \in \Re^p \mid n^T y \geq n^T y^0\}$ because of the strict convexity of C and the definition of the contingent cone. Hence $\text{Eff}(K(F,y^0)) = \text{Eff}(H) = \{y \in \Re^p \mid n^T y = n^T y^0\}$ holds. Thus we can prove a method for the determination of tradeoff directions for certain efficient outcomes by means of a generalization of Lyusternik's theorem in (Jahn 1996):

Theorem. Assume that F is closed and that there exists some neighbourhood U of $y^0 \in \text{bd}F$ in which the boundary of F coincides with some set $S = \{y \in \Re^p \mid h(y) = 0\}$, where $h: \Re^p \to \Re$ is differentiable on U and continuously differentiable in y^0 with $\nabla h(y^0) \neq 0$. If y^0 is an optimal solution of problem (P) and n is calculated according to equation (1), then

$$\text{Eff}(K(F,y^0)) = \{y \in \Re^p \mid n^T y = n^T y^0\}. \tag{2}$$

Finally, let us illustrate that problem (P) can be used to calculate tradeoffs of an efficient point even if the properly efficient point set is empty. Lee and Nakayama (1997) underlined by this example that the tradeoff concept of Henig and Buchanan (1997) can fail for non-convex feasible point sets in the outcome space.

Example.

Define $F := \{y \in \Re^2 \mid y_2 \geq 1\} \cup \{y \in \Re^2 \mid y_2 \geq -y_1\} \cup \{y \in \Re^2 \mid y_1 \geq 1\}$.

Then $G\text{-}Eff(F) = \emptyset$, but $y^0 := (0,0)^T \in Eff(F)$. y^0 solves (P) e.g. for parameters $w_1 = w_2 = b = 1$ and $v_1 = v_2 = -1$ with the optimal value $t_0 = 0$. Hence $Eff(K(F, y^0)) = \{y \in \Re^2 \mid n^T y = 0\}$ with $n = (1,1)^T$ is the set of tradeoff directions of F in y^0 in the sense of Lee and Nakayama.

References

Haimes YY, Chankong V (1979) Kuhn-Tucker multipliers as trade-offs in multiobjective decision making analysis. Automatica 15:59-72

Henig MI, Buchanan JT (1997) Tradeoff directions in multiobjective optimization problems. Mathematical Programming 78:357-374

Jahn J (1996) Introduction to the theory of nonlinear optimization. Springer-Verlag, Berlin Heidelberg New York

Kaliszewski I (1994) Quantitative Pareto analysis by cone separation technique. Kluwer Academic Publishers, Boston Dordrecht London

Kaliszewski I (2000) Using trade-off information in decision-making algorithms. Computers & Operations Research 27:161-182

Kaliszewski I, Michalowski W (1997) Efficient solutions and bounds on tradeoffs. J Optimization Theory Appl 94:381-394

Lee GM, Nakayama H (1997) Generalized trade-off directions in multiobjective optimization problems. Appl Math Lett 10:119-122

Miettinen K, Mäkelä MM (2002) On generalized trade-off directions in nonconvex multiobjective optimization. Math Progr Ser A 92:141-151

Pascoletti A, Serafini P (1984): Scalarizing vector optimization problems. J Optimization Theory Appl 42:499-524

Weidner P (1990) An approach to different scalarizations in vector optimization. Wissenschaftliche Zeitschrift der TH Ilmenau 36:103-110

Weidner P. (1993) Advantages of hyperbola efficiency. In: Brosowski B, Ester J, Helbig S, Nehse R (eds) Multicriteria decision. Peter Lang, Frankfurt aM, pp 123-137

Weidner P (1994) The influence of proper efficiency on optimal solutions of scalarizing problems in multicriteria optimization. OR Spektrum 16:255-260

Weidner P (2001) Problems in scalarizing multicriteria approaches. In: Köksalan M, Zionts S (eds) Multiple criteria decision making in the new millennium. Springer-Verlag, Berlin Heidelberg New York, pp 199-209

Yano H, Sakawa M (1987) Trade-off rates in the weighted Tchebycheff norm method. Large Scale Systems 13:167-177

A Soft Margin Algorithm controlling Tolerance Directly

Min Yoon[1], Hirotaka Nakayama[2], and Yeboon Yun[3]

[1] Yonsei University, 134 Shinchon-Dong Seodaemun-Gu, Seoul 120-749, Korea
[2] Konan University, 8-9-1 Okamoto, Higashinada, Kobe, Hyogo 658-8501, Japan
[3] Kagawa University, 2217-20 Hayashicho, Takamatsu, Kagawa 761-0396, Japan

Abstract. Generalization error bounds in Support Vector Machines are based on the minimum distance between training points and the separating hyperplane. The error of soft margin algorithm can be bounded by a target margin and some norms of the slack vector. In this paper, we propose a new method controlling allowable error and formulate considering the contamination by noise in data directly. The method can provide desirable separating hyperplanes easily by controlling a restricted slack parameter. Additionally, through an artificial numerical example, we compare the proposed method with a conventional soft margin algorithm.

1 Introduction

When considering large margin classifiers, where the complexity of a hypothesis is measured by its margin with respect to the data, the presence of noise leads to further problems. For example, solutions found by maximizing a margin are not stable with respect to the training points–slight modifications in the training set can significantly change the hypothesis– a brittleness which makes the maximal margin solution somehow undesirable. These problems have led to a method of "soft-margin", a procedure aimed at extending the large margin algorithms to noisy cases by permitting a slight sacrifice of accuracy. Using the existing soft margin algorithm, we can get a large margin, but the algorithms are very sensitive to penalty parameter in the formulated mathematical programming problem.

In this paper, we propose a new method controlling allowable error for noisy data. It will be shown that the proposed method can diminish the corruption of noise, and also make an improvement of generalization ability through a numerical example.

2 Error Bound for Soft Margin Algorithms

We begin the section by introducing the definition of margin slack variables[2,3], which show how much each point fails to meet the given threshold for classification. This threshold is called a target margin in this paper.

Definition 1. Let X be the space of data. For a real-valued classification function f on X and a target margin γ, if an example $(\boldsymbol{x}_i, y_i)) \in X \times \{-1, 1\}$

holds

$$y_i f(\boldsymbol{x}_i) \geqq \gamma,$$

then $\boldsymbol{x}_i$ is *correctly classified*: On the other hand, if $y_i f(\boldsymbol{x}_i) < \gamma$, then the $\boldsymbol{x}_i$ is *incorrectly classified.*

Definition 2. For a real-valued classification function f on X, the *margin slack variable* of an example $(\boldsymbol{x}_i, y_i) \in X \times \{-1, 1\}$ with respect to a function $f \in \mathcal{F}$, where $\mathcal{F}$is a class of real-valued functions, and a target margin γ is given by

$$\xi_i := \xi((\boldsymbol{x}_i, y_i), f, \gamma) = \max\{0,\ \gamma - y_i f(\boldsymbol{x}_i)\}.$$

For a training set S, the norm of ξ is given by

$$\|\xi\|_2. := \sqrt{\sum_{(\boldsymbol{x}_i, y_i) \in S} \xi((\boldsymbol{x}_i, y_i), f, \gamma)^2}$$

Note that $\xi_i > 0$ implies the incorrect classification of $(\boldsymbol{x}_i, y_i)$, because the points with non-zero $\xi((\boldsymbol{x}_i, y_i), f, \gamma)$ fail to achieve a positive margin of γ.

In terms of the target margin γ and 2-norm of the margin slack variables, the generalization error bound of soft margin algorithms for linear classification functions has been derived by Shawe-Taylor and Cristianini [3]:

Theorem 1. *Fix $\Delta > 0$. Consider a fixed but unknown probability distribution on the space $X \times \{-1, 1\}$ with support in the ball of radius R about the origin in X. Then with probability $1 - \delta$ over randomly drawn training sets S of size ℓ for all $\gamma > 0$, the generalization of a linear classifier $\mathbf{u}$ on X with $\|\mathbf{u}\| = 1$, thresholded at 0 is bounded by*

$$\epsilon(\ell, d, \delta) = \frac{2}{\ell}\left(d \log_2\left(\frac{8e\ell}{d}\right)\log_2(32\ell) + \log_2\left(\frac{8\ell}{\delta}\right)\right),$$

where

$$d = \left\lfloor \frac{64.5(R^2 + \Delta^2)(1 + \|\xi\|_2^2/\Delta^2)}{\gamma^2} \right\rfloor, \tag{1}$$

provided $\ell \geqq \frac{2}{\epsilon}$, $d \leqq e\ell$ and there is no discrete probability on misclassified training points.

Theorem 1 means that the generalization error is bounded by the amount how much the data fail to meet a target margin γ. The error bound is in terms of a norm of the slack variable, which implies that this quantity should be minimized in order to increase the generalization ability. The error bound does not rely on whether the training data are linearly separable or not, and hence can also handle the case when the data are corrupted by noise.

Furthermore, the error bound is monotonically increasing with respect to d, since for fixed ℓ and δ,

$$\begin{aligned}\left\{d\log_2\left(\frac{8e\ell}{d}\right)\right\}' &= \log_2 8e\ell - \log_2 d - \frac{1}{\log_e 2} \\ &> \log_2\left(\frac{8e\ell}{d}\right) - \frac{1}{\log_e 2} > 0. \ (\because\ d \leqq e\ell),\end{aligned}$$

where $'$ denotes the derivative.

In the next section, through a simple example, we will investigate the relation between d and a target margin γ in several linear classifier functions, and suggest a revised formulation of conventional soft margin algorithms.

3 The Proposed Method

Consider four cases of linear classifier functions f^1, f^2, f^3 and f^4 with 11 training data points as shown in **Fig. 1**. The classification function f^1 is perfectly separating the data, but the others f^2, f^3 and f^4 are not.

$$\begin{aligned} f^1(x_1,x_2) &= -x_1, & f^2(x_1,x_2) &= -4x_1 + 3x_2, \\ f^3(x_1,x_2) &= -3x_1 + 4x_2, & f^4(x_1,x_2) &= x_2 - 0.1. \end{aligned}$$

Without loss of generality, put $R = 1$ and $\Delta = 1$ in the equation (1). Then d for several target margins are shown in Table 1. We can find the fact that d is the smallest by (i) the linear classifier function f^1 for a target margin $\gamma \leqq 0.25$, (ii) f^2 for $0.30 \leqq \gamma \leqq 0.65$, (iii) f^3 for $0.7 \leqq \gamma \leqq 1.0$ and (iv) f^4 for $\gamma \geqq 1.05$, respectively.

For data without any noise, the classification may be well performed for a relatively small γ. In other words, completely separating hyperplanes such as f^1 may be the best for data without noise. For noisy cases, however, a relatively large γ is more desirable. How large γ is appropriate depends on

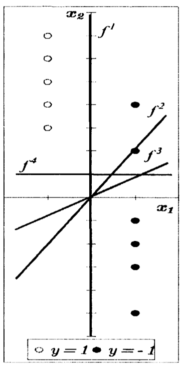


Fig. 1. example data and linear functions for classification

Table 1. Values of d for given a target margin γ

γ	d			
	f^1	f^2	f^3	f^4
0.05	51600	51740	52325	52684
0.10	12900	12998	13271	13416
0.15	5737	5833	6038	6135
0.20	3264	3342	3525	3597
0.25	2190	2210	2387	2445
0.30	1710	1625	1799	1851
0.35	1550	1312	1485	1532
0.40	1596	1158	1330	1375
0.45	1793	1118	1281	1326
0.50	2114	1171	1313	1361
0.55	2542	1306	1419	1467
0.60	3068	1520	1593	1638
0.65	3685	1810	1835	1875
0.70	4391	2175	2147	2182
0.75	5182	2613	2529	2559
0.80	6057	3125	2980	3005
0.85	7014	3711	3502	3522
0.90	8052	4371	4095	4108
0.95	9171	5105	4758	4766
1.00	10370	5914	5494	5495
1.05	11648	6797	6302	6296
1.10	13006	7756	7183	7170
1.15	14443	8790	8137	8117
1.20	15959	9900	9165	9137

how much we consider the influence of noise. In the soft margin algorithm, the slack variable $\xi_i,\ i = 1, \dots, \ell$ are introduced in order to take into account the influence of noise.

A conventional soft margin problem with ℓ_1-norm in SVMs can be formulated as follows[5]:

$$\begin{aligned} \underset{w, w_o, \xi_i}{\text{minimize}} \quad & ||\boldsymbol{w}||_1 + C\sum_{i=1}^{\ell} \xi_i \\ \text{subject to} \quad & y_i\left(\langle \boldsymbol{w}, \boldsymbol{x}_i\rangle + w_o\right) \geqq 1 - \xi_i, \\ & \xi_i \geqq 0, \ \ i = 1, \dots, \ell, \end{aligned} \tag{S}$$

where C is a weight parameter for slack variables and $\sum_{i=1}^{\ell} |w_i|$ is denoted by $||\boldsymbol{w}||_1$.

The optimal separating hyperplane can be found by solving the above problem (S) for the given parameter C. In practice, the parameter C is empirically chosen according to the influence of noise: for example, C is taken to be properly small if the influence of noise is large.

Using the soft margin algorithm (S), we find a separating hyperplane for the data employed in the previous section. In this case, let C=0.0001, 0.001, 0.01, 0.1, 1.0, 2.0. **Fig. 2** shows the optimal separating hyperplanes for several values of C. Although the parameter C is varying, we obtain only two kinds of hyperplanes. That is, we can not obtain non-vertical/non-horizontal hyperplanes such as f^2 and f^3 by the conventional soft margin

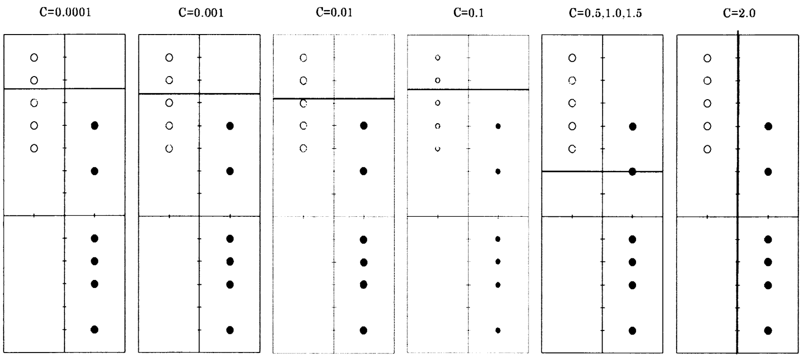


Fig. 2. separating hyperplanes by (S) varying C

algorithm whatever the value of C may be. This implies that we can not take into account the influence of noise in a subtle way.

In this paper, therefore, we suggest a new formulation in order to overcome these problems and control the influence of noise directly:

$$\begin{aligned} \underset{w, w_o, \xi_i}{\text{minimize}} \quad & \|w\|_1 + C \sum_{i=1}^{\ell} \xi_i \qquad (\mathrm{S}_{\xi_{\max}}) \\ \text{subject to} \quad & y_i \left(\langle w, x_i \rangle + w_o \right) \geqq 1 - \xi_i, \\ & 0 \leqq \xi_i \leqq \xi_{\max}, \;\; i = 1, \dots, \ell, \end{aligned}$$

where C is a weight parameter for slack variables and $\xi_{\max}$ is given as a fixed number.

In our formulation, the upper limit $\xi_{\max}$ to the margin slack variables, which means an allowable error, is introduced. We can control the influence of noise. For a relatively large value of $\xi_{\max}$, for example, we can take into account the influence of noise by $\xi_{\max}$ more directly than by the weight parameter C.

Fig. 3 shows the optimal separating hyperplanes obtained by the proposed method. For example, in the case of $C = 1.0$, compare the proposed method $(\mathrm{S}_{\xi_{\max}})$ with the existing method (S). The several forms of separating hyperplanes are produced by varying $\xi_{\max}$ in the proposed method , while the optimal separating hyperplane is only the horizontal by solving the problem (S). By adjusting the value of $\xi_{\max}$, it is possible to obtain an appropriate separating hyperplanes according to the degree how much the influence of noise should be considered. Since $\xi_{\max}$ itself plays a role as the level of allowable error, we can easily set the value of it.

4 Conclusion

In this paper, we have proposed a new formulation for the soft margin algorithm in order to consider the allowable error for noisy data. We have shown

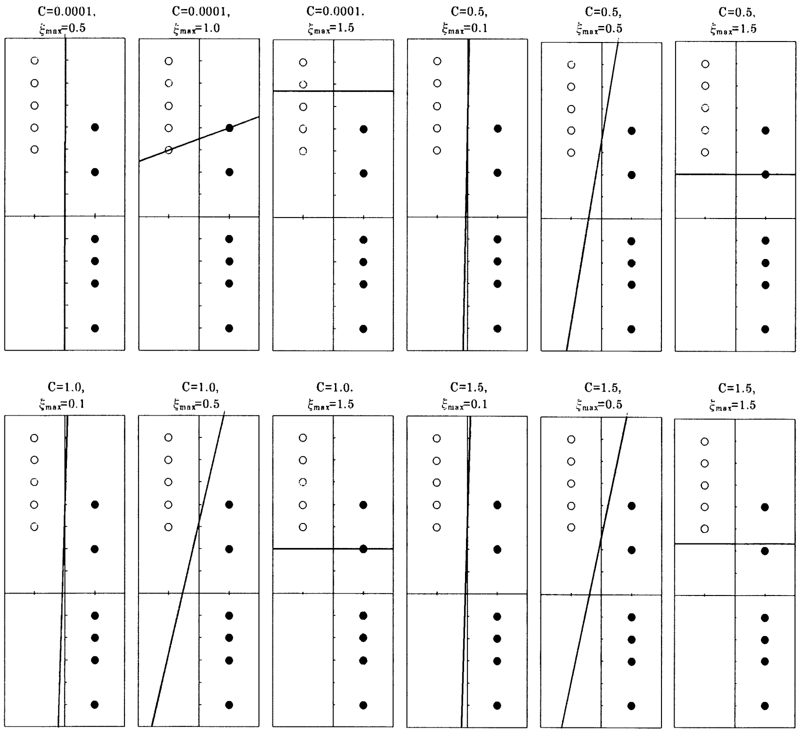


Fig. 3. separating hyperplanes varying C and $\xi_{\max}$ in the problem $(\mathrm{S}_{\xi_{\max}})$

that the proposed method can control the influence of noise directly and provide desirable separating hyperplanes easily by controlling a parameter $\xi_{\max}$. Therefore, it is expected that the proposed method can treat the corruption of noise effectively.

References

1. Bartlett, P. and Shawe-Taylor, J. (1999) Generalization Performance of Support Vector Machines and Other Pattern Classifiers, Advances in Kernel Methods-Support Vector Learning (edited by B. Schölkopf, C.J.C. Burges and A.J. Smola), MIT Press, 43–54
2. Cristianini, N. and Shawe-Taylor, J. (2000) An Introduction to Support Vector Machine, Cambridge University Press
3. Shawe-Taylor, J. and Cristianini, N. (2000) On the Generalisation of Soft Margin Algorithms, NeuroCOLT2 Technical Report Series, NC-TR-2000-082
4. Shawe-Taylor, J., Bartlett, P.L., Williamson, R.C.and M. Anthony (1998) Structural Risk Minimization over Data-Dependent Hierarchies, IEEE Transactions on Information Theory **44**(5), 1926–1940

5. Magasarian, O.L. (2000) Generalized Support Vector Machines, Advances in Large Margin Classifiers (edited by A. Smola, P. Bartlett, B. Shoolköpf and D. Schuurmans), 135–146.

An Analysis of Expected Utility based on Fuzzy Interval Data

Shin-ichi Yoshikawa [1], Tetsuji Okuda [2]

[1] Department of Management, College of Management, Nagoya Keizai University
[2] Department of Industrial Management, College of Engineering, Osaka Institute of Technology

Abstract. In this paper, we investigate an analysis of expected utility based on fuzzy interval data. We establish a data processing method that can treat fuzzy interval data. Unfortunately, the method that direct usage of membership functions of fuzzy interval data has the problems of the efficiency when we carry out a calculation. To solve such problems, we propose a practical method that uses the midpoints of membership functions as the representative values.

Keywords:Fuzzy Interval Data, F-moment, Expected Utility, Value of Fuzzy Information, Amount of Fuzzy Information

1. Introduction

The theory of expected utility is systematized in the field of statistical decision making. Conventional theory of expected utility is formulated based on probabilistic uncertainty. However, when there is human intervention, it is not easy to remove the data that includes vagueness.

This paper aims at giving the method of analysis of expected utility based on fuzzy interval data. Usually, the observation value is handled as a fixed value, but here it is regarded as the case of being able to obtain observation data whose boundaries of intervals are vague. These data are called fuzzy interval data, so it is desirable to establish a data processing method that is able to treat the vague data. However, the method that direct usage of membership functions of fuzzy interval data has the problems of the efficiency of calculations and so on. Therefore, a practical method that uses the midpoints of membership functions as the representative values is proposed in this paper. The problems mentioned above can be solved by our method.

2. Fuzzy Interval Data and Membership Functions

The vagueness is grasped as a fuzzy event (fuzzy set) μ_j defined on the interval $I_j = [x_j - h/2, x_j + h/2]$. In this paper, membership functions are defined on $[-h/2, h/2]$. They are represented by the following symmetrical function [1]. It is assumed as a basic form of fuzzy interval data.

〔trapezoidal type〕

$$\mu(x) = \max\left\{0, \min\left(1, -\frac{1}{2q}|x| + \frac{h}{8q} + \frac{1}{2}\right)\right\} \tag{1}$$

The q that appears on the right side of equation (1) is the shape parameter of the trapezoidal membership function and $0 < q \le h/4$. In order to statistically process fuzzy interval data μ_j, it is necessary to calculate P_j, the appearance probability of μ_j. From Zadeh's definition[2], this probability is expressed as

$$P_j = \int_{x_j - h/2}^{x_j + h/2} \mu_j(x) f(x) dx . \tag{2}$$

Here, defining F-moment of degree r of μ_j as

$$d_r = \int_{x_j - h/2}^{x_j + h/2} \mu_j(x)(x - x_j)^r dx , \tag{3}$$

it is possible to rewrite the equation (2) as

$$P_j = d_0 f(x_j) + \frac{d_2}{2!} f^{(2)}(x_j) + \frac{d_4}{4!} f^{(4)}(x_j) + \cdots + \frac{d_{2n}}{(2n)!} f^{(2n)}(x_j) + \cdots\cdots. \tag{4}$$

3. Expected Utility using Fuzzy Interval Data

We express the prior probability distribution of population parameter θ as $\xi(\theta)$ and the conditional density function as $f(x|\theta)$. Then we can introduce the posterior probability of θ by means of fuzzy interval data μ_j. As for the posterior probability of θ based on μ_j, using a Taylor expansion with the F-moment of μ_j, it becomes

$$\xi(\theta|\mu_j) = \xi(\theta|x_j) + \frac{d_2}{2d_0}\left\{\frac{f^{(2)}(x_j|\theta)\xi(\theta)}{f(x_j)} - \frac{\xi(\theta|x_j)}{f(x_j)}\int_{-\infty}^{\infty} f^{(2)}(x_j|\theta)\xi(\theta)d\theta\right\}$$
$$+ O(h^4). \tag{5}$$

In this way, the posterior probability $\xi(\theta \mid x_j)$ that uses the representative value x_j can be formulated by the usual Bayesian method, and by adding to it a compensation amount that is expressed in the second term on the right side of equation (5). Further, $O(h^4)$ in the equation (5) is the high order term. Usually, this term becomes a minute value. In the argument here, the width h of membership function is considered to the extent that it does not practically harm even if this term is ignored.

When we set $\xi(\theta)$ and $f(x \mid \theta)$ as the following function

$$\xi(\theta)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(\theta-\theta_0)^2}{2\sigma^2}\right\},\ f(x \mid \theta)=\frac{1}{\sqrt{2\pi}v}\exp\left\{-\frac{(x-\theta)^2}{2v^2}\right\} \quad (6)$$

respectively, the first term on the right side of equation (5) is expressed as

$$\xi(\theta \mid x_j)=\frac{1}{\sqrt{2\pi}\lambda}\exp\left\{-\frac{(\theta-\eta_j)^2}{2\lambda^2}\right\},$$

$$\eta_j=\left(\frac{\theta_0}{\sigma^2}+\frac{x_j}{v^2}\right)\left(\frac{1}{\sigma^2}+\frac{1}{v^2}\right)^{-1},\ \lambda=\left(\frac{1}{\sigma^2}+\frac{1}{v^2}\right)^{-1/2}. \quad (7)$$

And by setting the utility function to the action a_i as follows;

$$u(a_i,\theta)=K_i+k_i\theta \quad (8)$$

we can obtain the expected utility $U(a_i \mid \mu_j)$ written as

$$U(a_i \mid \mu_j)=K_i+k_i\eta_j+\frac{d_2}{d_0}\frac{k_i\lambda^2}{v^4}(\eta_j-x_j). \quad (9)$$

The first term and second term on the right side of equation (9) are described by the conventional method. When we treat the fuzzy interval data, we have to add the compensation amount that is expressed as the third term on the right side.

4.The Value of Fuzzy Information

We denote the best action as $a_{\mu j}$ which maximizes the expected utility. Then $U(a_{\mu j} \mid \mu_j)$ becomes as follows.

$$U(a_{\mu j} \mid \mu_j)=\max_{a_i}\{U(a_i \mid \mu_j)\} \quad (10)$$

The value of expected utility differs by the given fuzzy information μ_j, $j=-\infty,\cdots,\infty$, respectively. We consider the expected value with respect to μ_j. Then, the formulation of expected value $U(a_{\tilde{\mu}j} \mid \tilde{\mu}_j)$ is expressed as follows.

$$U(a_{\tilde{\mu}j} \mid \tilde{\mu}_j) = \sum_j U(a_{\mu j} \mid \mu_j)P(\mu_j)$$

$$= \sum_j \left[(K_{sj} + k_{sj}\eta_j) + \frac{d_2}{d_0}\frac{k_{sj}\lambda^2}{v^4}(\eta_j - x_j) \right] P(\mu_j) \qquad (11)$$

Here, we set the best utility function as

$$u(a_{\mu j},\theta) = K_{sj} + k_{sj}\theta . \qquad (12)$$

And when we can't get information at all, the best utility function is described as

$$U(a_0) = \int_{-\infty}^{\infty} u(a^0,\theta)\xi(\theta)d\theta$$

$$= K_0 + k_0\theta_0 . \qquad (13)$$

The a^0 in the equation (13) is the best action. Therefore, from equation (11) and (13), the value of fuzzy information is formulated as follows.

$$V(\tilde{\mu}_j) = U(a_{\tilde{\mu}j} \mid \tilde{\mu}_j) - U(a^0)$$

$$= \sum_j \left[(K_{sj} + k_{sj}\eta_j) + \frac{d_2}{d_0}\frac{k_{sj}\lambda^2}{v^4}(\eta_j - x_j) \right] P(\mu_j) - (K_0 + k_0\theta_0) \qquad (14)$$

5.The Amount of Fuzzy Information μ_j

The entropy on the state space $\{\theta\}$, when we can't get information at all, is formulated as

$$H(\xi) = -\int_{-\infty}^{\infty} \xi(\theta)\log\xi(\theta)d\theta . \qquad (15)$$

On the other hand, when we can get fuzzy information μ_j, the entropy is formulated as follows.

$$H(\xi(\mu_j)) = -\int_{-\infty}^{\infty} \xi(\theta \mid \mu_j)\log\xi(\theta \mid \mu_j)d\theta$$

$$= -\int_{-\infty}^{\infty} \xi(\theta \mid x_j)\log\xi(\theta \mid x_j)d\theta - \frac{d_2}{2d_0}\int_{-\infty}^{\infty} \{1 + \log\xi(\theta \mid x_j)\}$$

$$\times \left\{ \frac{f^{(2)}(x_j \mid \theta)\xi(\theta)}{f(x_j)} - \frac{\xi(\theta \mid x_j)}{f(x_j)}\int_{-\infty}^{\infty} f^{(2)}(x_j \mid \theta)\xi(\theta)d\theta \right\} d\theta \qquad (16)$$

Here, we consider the expected value with respect to μ_j. The expected value $M(\tilde{\mu}_j \mid \xi)$ is formulated as

$$M(\tilde{\mu}_j \mid \xi) = E_{\mu j}\{H(\xi(\mu_j))\}$$
$$= -\sum_j \left[\int_{-\infty}^{\infty} \xi(\theta \mid \mu_j) \log \xi(\theta \mid \mu_j)\right] P(\mu_j). \qquad (17)$$

Then, the amount of information $I(\tilde{\mu}_j \mid \xi)$ given by μ_j is expressed as

$$I(\tilde{\mu}_j \mid \xi) = M(\tilde{\mu}_j \mid \xi) - H(\xi). \qquad (18)$$

As for $I(\tilde{\mu}_j \mid \xi)$, we can show the following theorem.

[Theorem 1]

For any fuzzy information μ_j ($j = -\infty, \cdots, \infty$),

$$I(\tilde{\mu}_j \mid \xi) \geq 0 \qquad (19)$$

is assured.

The proof of **Theorem 1** is shown as follows.

$$I(\tilde{\mu}_j \mid \xi) = M(\tilde{\mu}_j \mid \xi) - H(\xi)$$
$$= \sum_j \left[\int_{-\infty}^{\infty} \xi(\theta \mid \mu_j) \log \xi(\theta \mid \mu_j)\right] P(\mu_j) - \sum_j \left[\int_{-\infty}^{\infty} \xi(\theta) P(\mu_j \mid \theta) \log \xi(\theta) d\theta\right]$$
$$= \sum_j \left[\int_{-\infty}^{\infty} \xi(\theta) P(\mu_j \mid \theta) \log \frac{\xi(\theta \mid \mu_j)}{\xi(\theta)} d\theta\right]$$
$$= \sum_j \left[\int_{-\infty}^{\infty} \xi(\theta \mid \mu_j) \log \frac{\xi(\theta \mid \mu_j)}{\xi(\theta)} d\theta\right] P(\mu_j)$$
$$\geq 0 \qquad (20)$$

6.Numerical Example

The **Table.1** shows a result of the value of fuzzy information and the amount of fuzzy information expressed by the right side of equation (14) and (18), respectively. In order to show the efficiency of our proposed method, the simulation here regards the handling of fuzzy interval data as the following three types (A), (B) and (C).

(A) Only the representative value
(B) Fuzzy interval data itself
(C) After compensation

As for (A), it is expressed by the central value of the membership function, and using only representative value, statistical processing is carried out. As for (B), it is a method of direct calculation and treating the membership function that does not depend on our proposed method. As for (C), this means the method that we are

Table.1. Result of the value of fuzzy information and the amount of fuzzy information

	(A)Only the representative value	(B)Fuzzy interval data itself	(C)After compensation
Equation (14)	1.814326	1.911372	1.904871
Equation (18)	0.353129	0.391658	0.391208

proposing here.

In our example, we assumed that $\xi(\theta)\sim N(80,20^2)$ and $f(x\,|\,\theta)\sim N(150,12^2)$, respectively. And as for the membership functions, we assumed that $h=10$ and $q=1.25$. Consequently, in regard to the value of fuzzy information and the amount of information in (A), they do not become accurate values based on fuzzy interval data respectively. But by compensating by means of (C), they become approximately the values derived from (B).

7.Conclusion

We have described a method of the analysis of expected utility based on fuzzy interval data, and this method's practicability has been clarified by the numerical example.

The method that only uses the representative values of membership functions, as does (A), it is not possible to carry out the proper processing of fuzzy interval data in which bias has occurred in the result. But the method (C) that is proposed here, the bias in (A) is removed. As for this method, the result becomes almost the same as that of the method (B) which handles fuzzy interval data in detail.

Consequently, with respect to actual problems from which vague data are obtained, the application of this method becomes possible.

References

[1]T.Okuda, Y.Kodono and K.Asai:Statistical Estimates based on Fuzzy Observation Data, Transactions of the Society of Instrument and Control Engineers, Vol.26, No.5, pp.564-571(1990).

[2]L.A.Zadeh:Probability Measures of Fuzzy Events, Journal of Mathematical Analysis and Applications, 28, pp.421-427(1965).

On Analyzing the Stability of Discrete Descriptor Systems via Generalized Lyapunov Equations

Qingling Zhang[1], James Lam[2], and Liqian Zhang[2]

[1] College of Science, Northeastern University, Shenyang, Liaoning 110004, CHINA
[2] Department of Mechanical Engineering, University of Hong Kong, Pokfulam Road, HONG KONG

Abstract. In this paper, the asymptotic stability of linear discrete time-invariant descriptor systems is studied via a generalized Lyapunov equation. The analysis covers both the causal and noncausal cases. In particular, the asymptotic stability of a discrete descriptor system (DDS) is related to the existence of a positive semidefinite solution of the generalized Lyapunov equation. The results strengthened those of earlier works for causal descriptor systems.

1 Introduction

It is well known that Lyapunov theory has played an important role in asymptotic stability analysis and related control problems. From the amount of literature on the subject, it is clear that Lyapunov equations for normal state space systems are well understood. However, only a few results related asymptotic stability to Lyapunov equations for descriptor systems [1–3]. Since descriptor systems arise in many engineering fields [1–4], it is very important to extend the well known Lyapunov theory to descriptor systems.

DDS may possess noncausal phenomenon which is alien to normal systems. The generalized Lyapunov equations for the systems, as proposed in previous works [5,6,4,7], cannot be directly utilized to analyze the asymptotic stability of noncausal DDS. These results are only valid for the causal case. The main issue here is to relate asymptotic stability for noncausal DDS to generalized Lyapunov equations. Zhang and Xu [4,7] investigated structural stability and linear quadratic control problems for causal DDS. Syrmos *et al.* [6] discussed the relationship between asymptotic stability and the solutions of a generalized Lyapunov equation for causal systems. Unfortunately, the noncausal discrete case has not been covered. Hence, it is very desirable to fill this gap and have a complete solution which provides a unification of the Lyapunov treatment to problems involving noncausal systems.

The aim of this paper is to analyze the asymptotic stability for both causal and noncausal DDS in terms of solutions of a generalized Lyapunov equation, which appeared in Zhang and Xu [7], as well as Syrmos *et al.* [6]. New conditions characterizing asymptotic stability are provided. In particular, the uniqueness of the positive semidefinite solution is discussed. It is also emphasized that the results obtained here are also valid for both causal DDS as well as normal discrete systems.

The outline of this paper is as follows. Related preliminary results are first presented in Section 2, then in Section 3 we study the solutions of a generalized Lyapunov equation and relate them to the asymptotic stability of DDS. For the limit of space, all proof is omitted.

2 Preliminaries

Consider a linear time-invariant DDS given by

$$Ex\left(k+1\right)=Ax\left(k\right)+Bu\left(k\right) \tag{1}$$

$$y\left(k\right)=Cx\left(k\right) \tag{2}$$

where $x\left(k\right)$, $u(k)$ and $y\left(k\right)$ are respectively the n-dimensional state vector, m-dimensional input vector and l-dimensional output vector; E, A, B and C are real matrices of appropriate dimensions. The DDS in (1) and (2) will be identified by its *realization* denoted by the quadruple (E, A, B, C) for short. In the sequel, whenever an argument, E, A, B, or C, of a realization is of no consequence in the development, we may replace it by a $*$. To ensure the existence and uniqueness of solution, DDS (E, A, B, C) is assumed to be *regular*, that is $\det(zE-A) \not\equiv 0$. The *finite poles* of DDS (E, A, B, C) are the finite roots of the characteristic equation $\det(zE-A)=0$. A DDS is said to be *asymptotically stable* (in Lyapunov sense) if all its finite poles lie in the open unit circle.

When the DDS is regular, there exist real invertible matrices P and Q such that

$$\bar{E}:=PEQ=\operatorname{diag}\left(I_{n_1},N\right),\quad \bar{A}:=PAQ=\operatorname{diag}\left(A_1,I_{n_2}\right) \tag{3}$$

where $A_1\in\mathbb{R}^{n_1\times n_1}$, $N\in\mathbb{R}^{n_2\times n_2}$ is nilpotent with nilpotent index h, i.e. $N^{h-1}\neq 0$, $N^h=0$, and $I_t\in\mathbb{R}^{t\times t}$ denotes the identity matrix of order t. Here, $n_1+n_2=n$. (3) is called the Weierstrass canonical form. The above transformation induces a conformal partition on B, C, and $x\left(k\right)$ such that

$$\bar{B}:=PB=\begin{bmatrix}B_1\\B_2\end{bmatrix},\quad \bar{C}:=CQ=\begin{bmatrix}C_1 & C_2\end{bmatrix},\quad Q^{-1}x=\begin{bmatrix}x_1\left(k\right)\\x_2\left(k\right)\end{bmatrix}. \tag{4}$$

Obviously, the DDS with realization $(\bar{E}, \bar{A}, \bar{B}, \bar{C})$, or

$$x_1\left(k+1\right)=A_1x_1\left(k\right)+B_1u\left(k\right),\quad y_1\left(k\right)=C_1x_1\left(k\right), \tag{5}$$

$$Nx_2\left(k+1\right)=x_2\left(k\right)+B_2u\left(k\right),\quad y_2\left(k\right)=C_2x_2\left(k\right), \tag{6}$$

$$y\left(k\right)=y_1\left(k\right)+y_2\left(k\right), \tag{7}$$

is *restricted system equivalent* to DDS (E, A, B, C). The above decomposition is generally referred to as a *forward-backward* decomposition. The solutions to (5) and (6) are given by

$$x_1(k) = A_1^t x_{10} + \sum_{i=0}^{k-1} A_1^{k-i-1} B_1 u(i) \ , \tag{8}$$

$$x_2(k) = -\sum_{i=0}^{h-1} N^i B_2 u(i+k) \ , \tag{9}$$

where $x_1(0) = x_{10}$, is a given admissible initial state vector.

Notice that $\det(zE - A) = \det(PQ)^{-1} \det(zI_{n_1} - A_1) \det(zN - I_{n_2})$ and $\det(zN - I_{n_2}) \neq 0$, thus DDS (E, A, B, C) is asymptotically stable if and only if the forward subsystem (5) is asymptotically stable. Although the asymptotic stability of DDS can be analyzed based on the above decomposition, Nichols pointed out that the decomposition is numerically sensitive to system data which may makes such decomposition sometimes unreliable [8]. Therefore, it is desirable to solve the problem without a decomposition.

DDS (E, A, B, C) is said to be *R-controllable* if its forward subsystem (5) is *controllable* which is equivalent to the fulfilment of the condition

$$\text{rank}\left(\begin{bmatrix} B_1 & A_1B_1 & \cdots & A_1^{n_1-1}B_1 \end{bmatrix}\right) = n_1.$$

The dual concept of *R-observability* can also be defined similarly and it can be established that *R-observability*, equivalent to the *observability* of the forward subsystem, of a DDS. Other equivalent conditions on these concepts may be found in [2].

The Laurent parameters ϕ_k, $-h \leq k < +\infty$, are introduced to specify the unique series expansion of the resolvent matrix about $z = \infty$,

$$(zE - A)^{-1} = z^{-1} \sum_{k=-h}^{\infty} \phi_k z^{-k} \quad .$$

Numerical reliable computation of ϕ_k can be found in [9]. In particular, in Weierstrass canonical form, we have [10]

$$\bar{\phi}_{-1} := Q^{-1} \phi_{-1} P^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & -I_{n_2} \end{bmatrix}.$$

Based on this, we have

$$E\phi_{-1} = P^{-1} \begin{bmatrix} I_{n_1} & 0 \\ 0 & N \end{bmatrix} Q^{-1} Q \begin{bmatrix} 0 & 0 \\ 0 & -I_{n_2} \end{bmatrix} P = P^{-1} \begin{bmatrix} 0 & 0 \\ 0 & -N \end{bmatrix} P \ ,$$

and for $W \in \mathbb{R}^{n \times n}$, $W = W^T$, we also have

$$\begin{aligned} \phi_{-1}^T E^T W E \phi_{-1} &= P^T \begin{bmatrix} 0 & 0 \\ 0 & -N^T \end{bmatrix} P^{-T} W P^{-1} \begin{bmatrix} 0 & 0 \\ 0 & -N \end{bmatrix} P \\ &= P^T \begin{bmatrix} 0 & 0 \\ 0 & N^T W_3 N \end{bmatrix} P \ , \end{aligned}$$

where

$$P^{-T}WP^{-1} = \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix}.$$

We define the set

$$\mathbb{W} = \left\{ W \in \mathbb{R}^{n\times n} \mid W \geq 0,\ \phi_{-1}^T E^T W E \phi_{-1} = 0 \right\}. \tag{10}$$

It is easily seen that $W \in \mathbb{W}$ if and only if $W \geq 0$ and

$$\phi_{-1}^T E^T W = 0 \quad \Leftrightarrow \quad N^T W_3 N = 0 \quad \Leftrightarrow \quad W_3 N = 0 .$$

Hence, $W \in \mathbb{W}$ is in the nullspace of matrix $\phi_{-1}^T E^T$. An admissible W can be obtained from the linear matrix equation $\phi_{-1}^T E^T W = 0$ under the restriction that $W \geq 0$.

3 Asymptotic Stability

To motivate the use of a generalized Lyapunov equation in the study of asymptotic stability of DDS, we note that

$$x_1(k) \neq 0 \quad \Rightarrow \quad Ex(k) \neq 0 \ ,$$

since

$$PEx(k) = PEQQ^{-1}x(k) = \begin{bmatrix} x_1(k) \\ Nx_2(k) \end{bmatrix} .$$

Thus one can construct a Lyapunov function of DDS (E, A, B, C) as

$$\mathcal{V}(Ex(k)) = x^T(k) E^T V E x(k) \ ,$$

where $V \in \mathbb{R}^{n\times n}$, $V \geq 0$, with the property that $\mathcal{V}(Ex(k)) > 0$, if $Ex(k) \neq 0$; $\mathcal{V}(0) = 0$, if $Ex(k) = 0$. Relating this function to DDS (1) and (2), we can see that $Ex(k) \neq 0$ is equivalent to $x_1(k) \neq 0$ when $x(k)$ is the solution (see (8) and (9) with $u(k) = 0$ and $t > 0$). The generalized Lyapunov equation associated to DDS (E, A, B, C) and $\mathcal{V}$ is given by

$$A^T V A - E^T V E = -E^T W E \ , \tag{11}$$

where $W \geq 0$. It is easily seen that (11) is a generalized form of Lyapunov equation in standard Lyapunov theory which corresponds to a discrete normal system if $E = I_n$. The generalized Lyapunov equation with $h = 1$ was considered in [11,6,7].

From (3), generalized Lyapunov equation (11) becomes

$$A_1^T V_1 A_1 - V_1 = -W_1 \ , \tag{12}$$

$$A_1^T V_2 - V_2 N = -W_2 N \ , \tag{13}$$

$$V_3 - N^T V_3 N = -N^T W_3 N \ , \tag{14}$$

where

$$P^{-T}VP^{-1} = \begin{bmatrix} V_1 & V_2 \\ V_2^T & V_3 \end{bmatrix} , \qquad P^{-T}WP^{-1} = \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix} , \tag{15}$$

such that the partitions are conformal to the dimensions of I_{n_1} and N. It is seen that for any given W_3, a closed-form solution for V_3 in (14) is given by

$$V_3 = -\sum_{i=1}^{h-1} \left(N^i\right)^T W_3 N^i \quad . \tag{16}$$

Moreover, from (16) we have $V_3 = 0$ if $W_3 \geq 0$ and $V_3 \geq 0$. We are now in a position to give a theorem which relates asymptotic stability of a DDS to its associated generalized Lyapunov equation.

Theorem 1. *DDS $(E, A, *, *)$ is asymptotically stable if and only if for any given $W \in \mathbb{W}$, which satisfies*

$$\operatorname{rank}\left(E^T W E\right) = \deg\det(zE - A) \ , \tag{17}$$

generalized Lyapunov equation (11) has unique solution $V \geq 0$ which satisfies

$$\operatorname{rank}(V) = \deg\det(zE - A) \ . \tag{18}$$

When $(E, A, *, *)$ is causal, i.e. $N = 0$, which is also equivalent to

$$\deg\det(zE - A) = \operatorname{rank}(E) , \tag{19}$$

we have the following result.

Corollary 1. *A causal DDS is asymptotically stable if and only if for any given $W > 0$, generalized Lyapunov equation (11) has unique solution $V \geq 0$ which satisfies* $\operatorname{rank}(V) = \operatorname{rank}(E)$.

Remark 1. Corollary 1 is related to Theorem 3.5 of [6] where in our case, strengthened to that solution V is unique. In that paper, it was only established that $E^T V E$ is unique. Theorem 1 gives a new asymptotic stability theorem which generalizes the previous results. From its proof, we see that generalized Lyapunov equation (11) is solvable if it is asymptotically stable. In general, the unique positive semidefinite solution V of generalized Lyapunov equation (11) may be expressed as

$$V \ = P^T \begin{bmatrix} V_1 & 0 \\ 0 & 0 \end{bmatrix} P.$$

In relating (11) to controllability and observability for DDS, we also have the following result.

Proposition 1. *Suppose DDS (E, A, B, C) is asymptotically stable. Then the DDS (E, A, B, C) with $C^TC \in \mathbb{W}$ (resp. $BB^T \in \mathbb{W}$) is R-observable (resp. R-controllable) if and only if the generalized Lyapunov equation*

$$A^TVA - E^TVE = -E^TC^TCE \tag{20}$$
$$(\text{resp.} AVA^T - EVE^T = -EBB^TE^T),$$

has unique solution $V \geq 0$, with

$$\text{rank}\,(V) = \deg\det(zE - A). \tag{21}$$

Remark 2. It should be mentioned that $C_2N = 0$ in (??) is equivalent to that the DDS is causal from state to output. This is seen from

$$y_2\,(k) = C_2x_2\,(k) = C_2Nx_2\,(k) = 0.$$

Meanwhile, if $BB^T \in \mathbb{W}$, then $B_2^TN^TNB_2 = 0$, which is equivalent to that the DDS is causal from input to state (see (9)). In general, from (9), we have

$$y_2\,(k) = C_2x_2\,(k) \;= -\sum_{i=0}^{h-1} C_2N^iB_2u\,(k+i) \;= -C_2B_2u(k),$$

and hence the DDS is causal from input to output if $C_2N = 0$ or $NB_2 = 0$. In practice, causality is essential for the implementation of a DDS.

References

1. Campbell, S. L. (1982): Singular Systems of Differential Equations. Pitman, London
2. Dai, L. (1989): Singular Control Systems. Spring-Verlag, Berlin
3. Lewis, F. L. (1986): A Survey of Linear Singular Systems. Circuits, Systems and Signal Processing. **5(1)**, 3–36
4. Zhang, Q. L. (1997): Decentralized and Robust Control for Large-scale Descriptor Systems. Northwestern Polytechnical University Press, Xian
5. Misra, P. (1992): Numerical Solution of Generalized Lyapunov Equation for Descriptor Systems. Presented at the Meeting of 30th Annual Allerton Conf. on Comm. Contr. Comp.
6. Syrmos, V. L. and P. Misra and R. Aripirala (1995): On the Discrete Generalized Lyapunov Equation. Automatica. **31(2)**, 297–301
7. Zhang, Q. L. and X. H. Xu (1994): Structural Stability and Linear Quadratic Control for Discrete Descriptor Systems. Proc. 1st Asian Contr. Conf. 407–410
8. Nichols, N. K. (1986): Robust Control System Design for Generalized State-space Systems. Proc. 25th IEEE Conf. Decision Contr. 538–540
9. Mertzios, B. G. and Lewis, F. L (1989): Fundamental Matrix of Discrete Singular Systems. Circuits Systems Signal Process. **8(3)**, 341–354
10. Bender, D. J. (1987): Lyapunov-like equations and reachability/observability grammians for descriptor systems. IEEE Trans on Automatic Control. **AC-32**, 343–348
11. Mertzios, B. G. (1984): On the Sensitivity Analysis of Linear Time-invariant Singular Systems. IEEE Trans. Cir. & Sys. **31**, 978-982

Solving DEA via Excel

Joe Zhu

Department of Management, Worcester Polytechnic Institute, Worcester, MA 01520 USA

Abstract The paper discusses how the Excel Solver can be used to solve various Data Envelopment Analysis (DEA) models. The structure of the DEA models allows one to use simple visual basic for applications (VBA) codes to automate the DEA calculation. As a result, one can easily establish the DEA spreadsheets and solve any existing or new DEA models.

Key words: data envelopment analysis (DEA), Excel, Solver.

1. Introduction

Data Envelopment Analysis (DEA) is designed to measure the relative efficiency within non-for-profit organizations where market prices are not available (Charnes et al. 1978). However, by its ability to model multiple-input and multiple-output relationships without *a priori* underlying functional form assumption, DEA has also been widely applied to other areas. The following DEA model is an input-oriented variable returns to scale (VRS) model where the inputs are minimized and the outputs are kept at their current levels.

$$\begin{aligned}
&\theta^* = \min \theta \qquad\qquad (1)\\
&\text{subject to}\\
&\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io} \qquad i = 1,2,\ldots,m;\\
&\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} \qquad r = 1,2,\ldots,s;\\
&\sum_{j=1}^{n} \lambda_j = 1\\
&\lambda_j \ge 0 \qquad j = 1,2,\ldots,n.
\end{aligned}$$

where DMU_o represents one of the n DMUs under evaluation, and x_{io} and y_{ro} are the ith input and rth output for DMU_o, respectively. If $\theta^* = 1$, then the current input levels cannot be reduced (proportionally), indicating that DMU_o is on the frontier. Otherwise, if $\theta^* < 1$, then DMU_o is dominated by the frontier. θ^* represents the (input-oriented) efficiency score of DMU_o.

By changing the constraint of $\sum_{j=1}^{n} \lambda_j$ in model (1), we can obtain other DEA models (see, e.g., Zhu (2002)). In the next section, we demonstrate how model (1) can be solved by Excel spreadsheets and the Excel Solver.

2. DEA Spreadsheets

We begin this section by organizing the data in Table 1 in a spreadsheet (see Figure 1). A spreadsheet model of (1) contains the following four major components: (i) cells for the decision variables (e.g., λ_j and θ); (ii) cell for the objective function (efficiency) (e.g., θ); (iii) cells containing formulas for computing the DEA reference set (the right-hand-side of the constraints) ($\sum_{j=1}^{n} \lambda_j x_{ij}, \sum_{j=1}^{n} \lambda_j y_{rj}$, and $\sum_{j=1}^{n} \lambda_j$); and (iv) cells containing formulas for computing the DMU under evaluation (left-hand-side of the constraints) (e.g., θx_{io} and y_{ro}).

Table 1. Data

Company	Assets	Equity	Employees	Revenue	Profit
Mitsubishi	91920.6	10950	36000	184365.2	346.2
Mitsui	68770.9	5553.9	80000	181518.7	314.8
Itochu	65708.9	4271.1	7182	169164.6	121.2
General Motors	217123.4	23345.5	709000	168828.6	6880.7
Sumitomo	50268.9	6681	6193	167530.7	210.5
Marubeni	71439.3	5239.1	6702	161057.4	156.6
Ford Motor	243283	24547	346990	137137	4139
Toyota Motor	106004.2	49691.6	146855	111052	2662.4
Exxon	91296	40436	82000	110009	6470
Royal Dutch/Shell Group	118011.6	58986.4	104000	109833.7	6904.6
Wal-Mart	37871	14762	675000	93627	2740
Hitachi	91620.9	29907.2	331852	84167.1	1468.8
Nippon Life Insurance	364762.5	2241.9	89690	83206.7	2426.6
Nippon Telegraph & Telephone	127077.3	42240.1	231400	81937.2	2209.1
AT&T	88884	17274	299300	79609	139

In Figure 1, cells I2 through I16 represent λ_j (j = 1, 2, ..., 15). Cell F19 represents the efficiency score θ which is the objective function.

For the DEA reference set (left-hand-side of the envelopment model), we enter the following formulas that calculate the weighted sums of inputs and outputs across all DMUs, respectively.

Cell B20 =SUMPRODUCT(B2:B16,$I$2:$I$16)
Cell B21 =SUMPRODUCT(C2:C16,$I$2:$I$16)
Cell B22 =SUMPRODUCT(D2:D16,$I$2:$I$16)
Cell B23 =SUMPRODUCT(F2:F16,$I$2:$I$16)

Cell B24 =SUMPRODUCT(G2:G16,$I$2:$I$16)

For the DMU under evaluation (DMU1: Mitsubishi), we enter the following formulas into cells D21:D24.

Cell D20 =$F$19*INDEX(B2:B16,E18,1)
Cell D21 =$F$19*INDEX(C2:C16,E18,1)
Cell D22 =$F$19*INDEX(C2:C16,E18,1)
Cell D23 =INDEX(F2:F16,E18,1)
Cell D24 =INDEX(G2:G16,E18,1)

	A	B	C	D	E	F	G	H	I	J
1	Company	Assets	Equity	Employees		Revenue	Profit		λ	changing cell
2	Mitsubishi	91920.6	10950	36000		184365.2	346.2		1	changing cell
3	Mitsui	68770.9	5553.9	80000		181518.7	314.8		0	changing cell
4	Itochu	65708.9	4271.1	7182		169164.6	121.2		0	changing cell
5	General Motors	217123.4	23345.5	709000		168828.6	6880.7		0	changing cell
6	Sumitomo	50268.9	6681	6193		167530.7	210.5		0	changing cell
7	Marubeni	71439.3	5239.1	6702		161057.4	156.6		0	changing cell
8	Ford Motor	243283	24547	346990		137137	4139		0	changing cell
9	Toyota Motor	106004.2	49691.6	146855		111052	2662.4		0	changing cell
10	Exxon	91296	40436	82000		110009	6470		0	changing cell
11	Royal Dutch/Shell Group	118011.6	58986.4	104000		109833.7	6904.6		0	changing cell
12	Wal-Mart	37871	14762	675000		93627	2740		0	changing cell
13	Hitachi	91620.9	29907.2	331852		84167.1	1468.8		0	changing cell
14	Nippon Life Insurance	364762.5	2241.9	89690		83206.7	2426.6		0	changing cell
15	Nippon Telegraph & Telephone	127077.3	42240.1	231400		81937.2	2209.1		0	changing cell
16	AT&T	88884	17274	299300		79609	139		0	
17							Reserved to indicate the DMU under evaluation.			
18		Reference		DMU under	1	Efficiency				
19	**Constraints**	set		Evaluation		1				
20	Assets	91920.6	≤	91920.6						
21	Equity	10950	≤	10950		Efficiency; θ; A changing cell; Target cell in Solver				
22	Employees	36000	≤	36000						
23	Revenue Constraint on Σλ	184365.2	≥	184365.2						
24	Profit	346.2	≥	346.2						
25	Σλ	1	=	1						

Fig. 1. DEA Spreadsheet Model

Finally, we enter the formula for $\sum_{j=1}^{n} \lambda_j = 1$ into cells B25 (=SUM(I2:I16)) and D25 (=1), respectively.

Cell E18 is reserved to indicate the DMU under evaluation. The function INDEX(array,row number,column number) returns the value in the specified row and column of the given array. Because cell E18 contains the current value of 1, the INDEX function in cell D23 returns the value in first row and first column of the Revenue array F2:F16 (or the value in cell F2, the Revenue output for DMU1). When the value in cell E18 changes from 1 to 15, the INDEX functions in cells D20:D24 return the input and output values for a specific DMU under evaluation. This feature becomes obvious and useful when we provide the Visual Basic for Applications (VBA) code to automate the DEA computation.

After the DEA model is set up in the spreadsheet, we can use Solver to find the optimal solutions. First, we need to invoke Solver in Excel by using the Tools/Solver menu item. Now, you should see the Solver Parameters dialog box shown in Figure 2

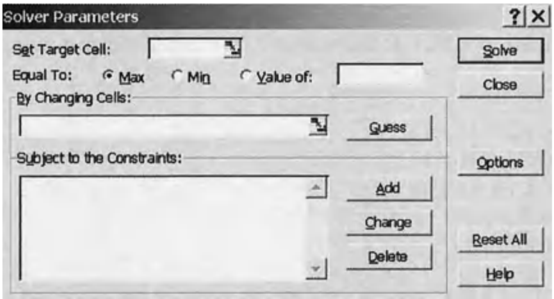


Fig. 2. Excel Solver Parameters

Set Target Cell indicates the objective function cell in the spreadsheet, and whether its value should be maximized or minimized. In our case, the target cell is the DEA efficiency represented by cell F19, and its value should be minimized, because of model (1). If the DEA model is output-oriented, then choose max.

Changing Cells represent the decision variables in the spreadsheet. In our case, they represent the λ_j (j = 1,2, ..., 15) and θ, and should be cells I2:I16 and F19, respectively

Constraints represent the constraints in the spreadsheet. In our case, they are determined by cells B20:B25 and D20:D25. For example, click the Add button shown in Figure 2, you will see the Add Constraint dialog box shown in Figure 3.

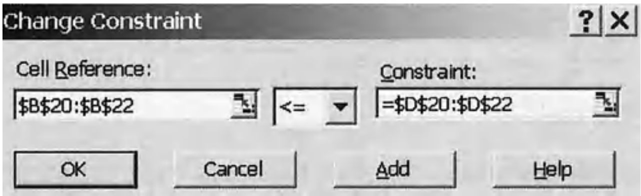


Fig. 3. Adding Constraints

In the spreadsheet model shown in Figure 1, we have six constraints. The "Cell Reference" corresponds to the DEA Reference Set, and "Constraint" corresponds to the DMU under evaluation. The first three constraints are related to the three inputs (see Figure 3). Click the Add button to add additional constraints (output constraints and $\sum_{j=1}^{n} \lambda_j = 1$), and click the OK button when you have finished adding the constraints. The set of the constraints are shown in Figure 4. Note that λ_j and θ are all non-negative, and the model (1) is a linear programming problem. This can be achieved by clicking the Option button in Figure 2, and then checking the Assume Non-Negative and Assume Linear Model boxes. Now, we have successfully set up the Solver Parameters dialog box, as shown in Figure 4. Click the Solve button to solve the model.

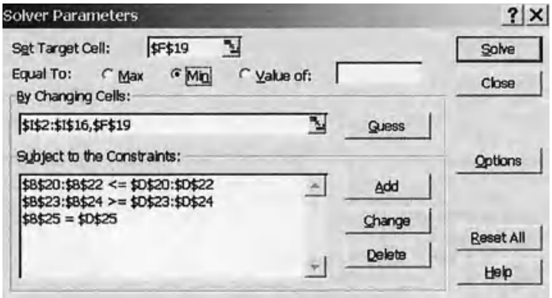


Fig. 4. Solver Parameters for DEA Model

To complete the analysis for the remaining 14 companies, one needs to manually change the value in cell E18 to 2, 3, ..., 15 and use Solver to re-optimize the spreadsheet model for each company and record the efficiency scores (in column J, for instance). When the number of DMUs becomes large, the manual process is apparently cumbersome.

Note that exactly the same Solver settings will be used to find the optimal solutions for the remaining DMUs. This allows us to write a simple VBA code to carry out the process automatically.

Before we write the VBA code, we need to set a reference to Solver Add-In in Visual Basic (VB) Editor. Otherwise, VBA will not recognize the Solver functions and you will get a "Sub or function not defined" error message.

We may follow the following procedure to set the reference. Enter the VB Editor by pressing *Alt-F11* key combination (or using the Tools/Macro/Visual Basic Editor menu item). Open the Tools/References menu in the VB Editor. This brings up a list of references. One of these should be **Solver.xla** To add the reference, simply check its box.

After the Solver reference is added, we should see "Reference to Solver.xla" under the "References" in the VBA Project Explorer window. Next, select the Insert/Module menu item in the VB Editor. This action will add a Module (e.g., Module1) into the Excel file.

Now, we can insert the VBA code into the Module1. Type "Sub DEA()" in the code window. This generates a VBA procedure called DEA which is also the Macro name. Figure 5 shows the VBA code for automating the DEA calculation.

The Macro statement "SolverSolve UserFinish:=True" tells the Solver to solve the DEA problem without displaying the Solver Results dialog box. The "Offset(*rowOffset*, *columnOffset*)" property takes two arguments that correspond to the relative position from the upper-left cell of the specified Range. When we evaluate the first DMU, i.e., DMUNo = 1, Range("J1").Offset(1,0) refers to cell J2. The statements "With Range("J1") and ".Offset(DMUNo, 0)=Range("F19") take the optimal objective function value (efficiency score) in cell F19 and place it in cell J "DMUNo" (that is, cell J2, J3, ..., J16).

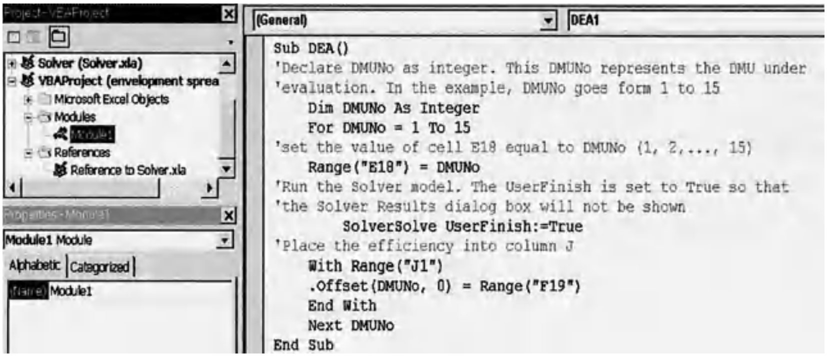


Fig. 5. VBA Code for DEA

Enter the Run Macro dialog box by pressing *Alt-F8* key combination (or using the Tools/Macro/Macros menu item). You should see "DEA". Select "DEA" and then click the Run button. This action will generate the efficiency scores.

3. Conclusions

This paper shows how DEA models can be solved in Excel spreadsheets. Although the paper only deals with one DEA model, the procedure can be applied to other DEA models. Zhu (2002) provides a detailed discussion on solving DEA via Excel spreadsheets and Excel Solver. Zhu (2002) also provides an easy-to-use DEA software — DEA Excel Solver. This DEA Excel Solver is an Add-In for Microsoft Excel and provides a custom menu of DEA approaches which include more than 150 different DEA models. It is an extremely powerful tool that can assist decision-makers in benchmarking and analyzing complex operational efficiency issues in manufacturing organizations as well as evaluating processes in banking, retail, franchising, health care, e-business, public services and many other industries. The DEA Excel Solver does not set limit on the number of units, inputs or outputs. With the capacity of Excel Solver, the DEA Excel Solver can deal with large sized performance evaluation tasks.

References

Charnes, A., Cooper, W.W., and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research* 2/6, 429-444.

Zhu, Joe (2002), Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets. Kluwer Academic Publishers, Boston.

PART III:

General Papers – Applications

Planning and Scheduling Staff Duties by Goal Programming

Sydney CK Chu[1] and Christina SY Yuen[2]

[1] Univ of Hong Kong
[2] Hongkong Airport Services

Abstract. We propose goal programming (GP) models for an integrated problem of staff duties planning and scheduling, for baggage services section staff at the Hong Kong International Airport. The problem is solved via its decomposition into a GP planner, followed by a GP scheduler. The results can be adopted as a good crew schedule in the sense that it is both feasible, satisfying various work conditions, and "optimal" in minimizing overtime shifts.

1 Introduction

This paper advocates a general modeling framework for a complete crew assignment system. It arises naturally as a mathematical description for the staff deployment problem of their baggage handling agents at BSS-HAS, the Baggage Services Section of the Hongkong Airport Services, Ltd. HAS of the (new) Hong Kong International Airport (at Chak Lap Kok of Lantau Island) is the primary handler of all ground services and support functions, including aircrafts and passengers alike.

Our project of optimization modeling for staffing is motivated by the need to produce daily work plan of the baggage service agents at the passenger terminal. Our complete BSS crew system consists of its three component GP models: the Duties Generation Problem (DGP), the Crew Scheduling Problem (CSP) and the Crew Rostering Problem (CRP). While such modeling may well be regarded as one among the vast literature of the commonly known area of workforce planning/scheduling (an excellent review is given by Bodin et al, 1983), our decomposition approach has, for the actual case study, exhibited its significant impact albeit its modeling simplicity. The resulting preemptive goal programming formulations have very satisfactorily addressed the planning/scheduling/rostering issues to handle frequent changes of flight schedules by flexibility in work patterns of agent duties.

1.1 Crew Scheduling

In the general area of routing and scheduling of vehicle and crew (Bodin et al, 1983), it is common to separate the overall problem into two steps consisting of the determination of the time tables – vehicle routing, followed by the staff assignment – crew scheduling.

Various useful models for crew scheduling problem (CSP) aiming at differing merits and purposes have been proposed, such as (matching based) heuristics models of Ball et al, 1983; network models of Carraresi and Gallo, 1984; and set partitioning models of Falkner and Ryan, 1987. Among the mathematical programming approaches, there are work of Lessard et al, 1981; column generation approach of Desrochers and Soumis, 1989, Desrochers et al, 1992; integer programming approach of Ryan and Foster, 1981, Ryan and Falkner, 1987; decomposition approaches of Patrikalakis and Xerocostas, 1992, Vance et al, 1997; and complementary approaches of Wren et al, 1985.

These quoted above constitute only a tiny fraction of the vast literature, not to mention techniques of implementation for practical applications, notably computerized scheduling such as the various reported systems of "HASTUS" of Lessard et al, 1981, "CREW-OPT" by Desrochers, et al, 1992, "EXPRESS" by Falkner and Ryan, 1992; and that of Chu and Chan, 1998.

Successful real applications are extremely significant for the airlines. Besides the "household name" of SABRE, we mention two most recent "milestone" works of Vance et al, 1997 and of Mason et al, 1998.

1.2 Crew Rostering

The outcome of the crew scheduling phase is typically a set of daily staff assignments required to cover the (actual or forecast) demand. "In the (next) *crew rostering* phase, a set of *working rosters* is constructed that determine the sequence of duties that each single crew has to perform ..., to cover everyday all the duties selected in the first phase" (quoted from Caprara et al, 1998). This has been referred to as the Crew Rostering Problem (CRP) by Caprara et al in their FARO prize winning work for the Italian Railway *Ferrovie dello Stato SpA*, jointly sponsored by the Italian Operational Research Society during 1994-1995.

Similar to the case of crew scheduling, past work on CRP has seen numerous approaches and applications. There are optimization approaches such as that of Gamache and Soumis, 1993; network model of Balakrishnan and Wong, 1990; and column generation approach of Gamache et al, 1994. Novel heuristics approaches integrating Set Covering and/or Assignment Problem are reported by Hagberg, 1985, Carpaneto and Toth, 1987, and Caprara et al, 1995. More recently, Valouxis and Housos, 2002, propose a quick heuristics for combined bus and driver scheduling, consisting of minimum cost matching, set partitioning and shortest path.

1.3 Duties Generation

The modeling formulation of DGP that we put forth here can be interpreted as the basic core – the planner – of a more sophisticated DGP/CSP/CRP integrated model in the following sense. DGP in its simplest form (computes and) allocates duties (of given fixed structure of work pattern, rather than crew or staff needing further varying requirements of scheduling) to cover known demands. Demands are given, for equally spaced (such as half-hourly) time intervals of (the working time of) a day. As such, DGP is the *prerequisite* to CSP and CRP in that it provides the planning inputs needed in subsequent scheduling and rostering of staff. The logical flow of their relationships can be summarized below, where t = hour of day, j = day of week, p = weekly work pattern and a = agent,

Hourly Demands $D(t)$
<<Planner: DGP Model>>
Duties $X(i,k)$
Allocations $R(j,t)$
↓
Daily Requirements $R(j) = \sum_t R(j,t)$
<<Scheduler: CSP Model>>
Daily Staffing $S(j)$
Allocations $R(j,t)$
↓
<<Roster: CRP Model>>
Duty Rosters $I(p,a,j,t)$

2 Goal Programming Models

As its name implies, DGP allocates duties (performed by crew) in an optimal way to meet known demand over a contiguous number of time intervals. We describe only its *extended* formulation below. A detailed account of DGP formulations is given in an earlier paper of Chu, 2001.

DGP Model
We use the following common notations for all the subsequent models. Let H be the working time horizon, and $h = 1, \cdots, H$ index the individual hours (or half-hours). R_h denotes the demand for interval h and d_h represents the over allocation (or over-achievement deviation variable in a goal programming context) at interval h.

The length of a duty is denoted by J. The primary decision variable x_{ij} is the number of allocated staff that starts duty from interval i and breaks at the j^{th} interval after the start of duty, $j = 1, \cdots, J$. Hence for a working horizon of intervals $1 \cdots H$, we have for the index $i = S, \cdots, T$. The earliest start interval S is such that $S \geq 1$ whereas the latest start interval T is limited to $T \leq H - J + 1$ (to finish work at interval H). Note that normally $S = 1$ as long as $R_1 > 0$ (there is demand for the very first interval); and $T = H - J + 1$ whenever $R_H > 0$ (there is demand for the very last interval).

As noted by Mason et al, 1998, personnel scheduling problems (or referred to as *workforce allocation* problems by Baker, 1976) have been studied for many years. Network flow formulations, such as in Segal, 1974 and Bartholdi and Ratliff, 1978 can well handle their simplest forms. Additional side constraints such as break requirements demand more complex procedures.

One advantage of the DGP model is its ease of extension in various ways. One such concern is the inclusion of flexibility in staffing mode: introducing over-time (OT) for any number of on-duty staff. This simply calls for adding another decision variable y_{mn} representing the number of allocated OT staff who start work at interval m and finish work in interval n. Defined generally as such, OT work can take different modes: a (limited) number of intervals immediately before a regular time (RT) duty only, or a (similarly limited) number of intervals immediately after an RT duty only.

As an illustration, the DGP model with OT allocation is given below.

$$\text{Min} \quad \sum_{m=1}^{H} \sum_{n=m}^{m+L-1} \Big(\sum_{h=m}^{n} T_h\Big)\, y_{mn} \tag{1a}$$

$$\text{Min} \quad \sum_{i=S}^{T} \sum_{j=1}^{J} c_{ij} x_{ij} \tag{1b}$$

$$\text{Min} \quad WD \tag{1c}$$

Subject to

$$\sum_{i=p}^{q} \sum_{j \neq h-i+1} x_{ij} + \sum_{m=h-L+1}^{h} \sum_{n=h}^{m+L-1} y_{mn} - d_h = R_h\,, \quad h = 1, \cdots, H \tag{2}$$

$$\sum_{m=(i-1)-L+1}^{i-1} y_{m,i-1} + \sum_{n=i+J}^{(i+J)+L-1} y_{i+J,n} \le \sum_{j=1}^{J} x_{ij}\,, \quad i = S, \cdots, T \tag{3}$$

$$\sum_{i=S}^{T} \sum_{j=1}^{J} x_{ij} \le \text{MaxRT} \tag{4}$$

$$\sum_{m=1}^{H} \sum_{n=1}^{H} y_{mn} \le \text{MaxOT} \tag{5}$$

$$d_h \le D\,, \qquad h = 1, \cdots, H \tag{6}$$

Here $p \equiv \max\,\{h - J + 1, S\}$, $q \equiv \min\,\{h, T\}$, and both the RT allocation $\{\ x_{ij}\ \}$ and the OT allocation $\{\ y_{mn}\ \}$ are non-negative integer variables. Note that L stands for the maximum number of (additional) OT intervals allowed before or after the J RT intervals.

We see that the LHS of constraint (2) is the total work contribution as a function of both RT and OT staff. The RT (or x_{ij}) portion is straight-forward, while the OT (or y_{mn}) part picks out the total number of OT staff for a maximum span of L intervals which cover h. Constraint (3) ensures that each

y_{mn} is indeed an OT allocation, by stipulating that an OT is assigned only if there is already an RT x_{ij} allocation (before or after). The single parameter *MaxRT* of constraint (4) denotes the maximum number (or strength) of RT staff and that *MaxOT* of constraint (5) is the maximum permitted number of OT staff. Suitably mixed (i.e. RT+OT) duties allocation can be obtained by varying these two parameters in repeated runs of the model, preemptively with the three goals (1a) to (1c) in that order.

Finally, the coefficients $\{ T_h \}$ in (1a) represent the unit OT pay rates, possibly different for different time intervals of the day, whereas the coefficients $\{ c_{ij} \}$ in (1b) represent the usual unit RT pay rates. The single variable D of constraint (6) records the maximum (i.e. over achievement) deviation over all time intervals. Its non-smoothed penalty term WD in (1c) is treated as a lowest priority goal in search of an "optimal" mixed duties allocation plan that includes both (positive) x_{ij} and y_{mn}.

CSP Model

Next, the CSP model, which is often referred to as the (cyclic or weekly) staffing model (see, for example, Schrage, 1999) is stated below.

Inputs: (from DGP Model)

$R(j) \equiv \sum_t R(j,t)$ Required no. of start duties on day j

Constratints:

$$\sum_{1 \le i \le 5} \text{START}(@wrap(j-i+1,7)) - \text{OVER}(j) = R(j),\ j = 1,\dots,7$$

Objective functions:

$$\text{Min} \sum_i \text{START}(i)$$

$$\text{Min}\ \ \text{MaxOVER}\ \ (\equiv\ \text{Max}_j\ \text{OVER}(j))$$

CRP Model

Finally, the CRP model, which in many ways has the interpretation of a set-covering formulation is given below.

Indices:

$p =$ roster pattern $(1,\dots,7)$
$a =$ baggage service agent (BSA)
$j =$ start time half-hour $(1,\dots,11$, or $12,\dots,22)$
$t =$ day of week $(1,\dots,7)$

Inputs:

$R(j,t) =$ required no. of start duties (at half-hour j on day t)
— output from DGP Model
$S(t) =$ required no. of starting crew (on day t)
— output from CSP Model

Covering (Roster) Variables:

$I(p,a,j,t) = 1,$
if agent a is assigned to cover roster pattern p at time j on day t
$I(p,a,j,t) = 0$, if one or more of the following conditions hold:

i) $a > S(p)$
ii) $t = @\ \text{warp}(p+5,7)\ [\equiv t1(p)]$
iii) $t = @\ \text{warp}(p+6,7)\ [\equiv t2(p)]$
iv) $R(j,t) = 0$

Covering (Roster) Constraints:

1) Each (assigned) agent gets 1 duty on each working day

$$\sum_{j|R(j,t)\geq 1} I(p,a,j,t) = 1, \quad \forall p, a \leq S(p), t \neq t1(p), t2(p)$$

2) Each (assigned) agent gets 5 duties each week

$$\sum_{j,t|R(j,t)\geq 1, t\neq t1(p),t2(p)} I(p,a,j,t) = 5, \quad \forall p, a \leq S(p)$$

3) Start duties ($R(j,t)$) of each slot are covered

$$\sum_{p,a\leq S(p)} I(p,a,j,t) \geq R(j,t), \quad \forall j,t|R(j,t) \geq 1; t \neq t1(p), t2(p)$$

4) Start rosters ($S(t)$) of each day are allocated

$$\sum_{p,a\leq S(p)} \sum_{j|R(j,t)\geq 1} I(p,a,j,t) - D(t) = S(t), \quad \forall t \neq t1(p), t2(p)$$

where $D(\cdot)$ is the over allocation to be minimized in the objective function.

3 A Concluding Remark

The purpose of this paper is to illustrate by way of this DGP/CSP/CRP modeling and computational experience, the advantage of its readily producing significant improvement over existing manual staff assignment. Its usefulness is somehow, in our opinion and experience of actually applying it in real situations, rather highly out of proportion with regard to its modeling simplicity. The system's usefulness to the HAS users is indeed decreasing from planning (DGP), to scheduling (CSP), and finally to dispatching (CRP). The last is still influenced regularly by day-to-day actual dispatching and rostering needs (which are left more to the field operational supervisors).

Acknowledgement — This work was initiated by the Management of HAS (the Hong Kong Airport Services, Limited) among other operational projects of similar logistics nature we have been conducting at the Hong Kong International Airport. Deep appreciation is expressed for their provision of information and data, as well as numerous useful discussions.

References

1. Baker, K. (1976). Workforce allocation in cyclical scheduling problems: a survey. *Operational Research Quarterly*, Vol. 27, pp. 155-167
2. Balakrishnan, N. & Wong, R.T. (1990). A network model for the rotating workforce scheduling problem. *Networks*, Vol. 20, pp. 25-42
3. Ball, M., Bodin, L. & Dial, R. (1983). A matching based heuristics for scheduling mass transit crews and vehicles. *Transportation Science*, Vol. 17, pp. 4-31
4. Bartholdi, J. & Ratliff, H. (1978). Unnetworks, with applications to idle time scheduling. *Management Science*, Vol. 4, pp. 850-858
5. Bodin, L., Golden, B., Assad, A & Ball, M. (1983). Routing and scheduling of vehicles and crews: the state of the art. *Computer and Operations Research*, Vol. 10, pp. 63-211
6. Caprara, A., Fischetti, M. & Toth, P. (1995). A heuristic method for the set covering problem. Technical Report, DEIS OR-95-8, University of Bologna.
7. Caprara, A., Toth, P., Vigo, D. & Fischetti, M. (1998). Modeling and solving the crew rostering problem. *Operations Research*, Vol. 46, pp. 820-830
8. Carpaneto, G. and Toth, P. (1987). Primal-dual algorithms for the assignment problem. *Discrete Applied Mathematics*, Vol. 18, pp. 137-153
9. Carraresi, P. & Gallo, G. (1984). Network models for vehicle and crew scheduling. *European Journal of Operational Research*, Vol. 16, pp. 139-151
10. Chu, S.C.K. (2001). A goal programming model for crew duties generation. *Journal of Multi-criteria Decision Analysis*, Vol. 10, pp.143-151
11. Chu, S.C.K. & Chan, E.C.H. (1998). Crew scheduling of light rail transit in Hong Kong: from modeling to implementation. *Computers & Operations Research*, Vol. 25, pp.887-894
12. Desrochers, M., Gilbert, J., Sauve, M. & Soumis, F. (1992). CREW-OPT: Subproblem modeling in a column generation approach to the urban transit crew scheduling problem. In M. Desrochers & J.M. Rousseau (Eds.) *Computer-Aided Transit Scheduling*, Lecture Notes in Economics and Mathematical Systems 386. Springer-Verlag, pp.395-406
13. Desrochers, M. & Soumis, F. (1989). A column generation approach to the urban transit crew scheduling problem. *Transportation Science*, Vol. 23, pp. 1-13
14. Falkner, J.C. and Ryan, D.M. (1987). Aspects of bus crew scheduling using a set partitioning model. *Computer-Aided Transit Scheduling: Proceedings of the Fourth Conference*. Springer-Verlag, pp. 91-103
15. Falkner, J.C. & Ryan, D.M. (1992). EXPRESS: Set partitioning for bus crew scheduling in Christchurch. In M. Desrochers & J.M. Rousseau (Eds) *Computer-Aided Transit Scheduling*, Lecture Notes in Economics and Mathematical Systems 386. Springer-Verlag, pp.359-378

16. Gamache, M. & Soumis, F. (1993). A method for optimally solving the rostering problem. Les Cahiers du GERAD, G-93-40. Montréal
17. Gamache, M., Soumis, F., Marquis, G. & Desrosiers, J. (1994). A column generation approach for large scale aircrew rostering problems. Les Cahiers du GERAD, G-94-20. Montréal
18. Hagberg, B. (1985). An assignment approach to the rostering problem. In J.M. Rousseau (Ed.) *Computer Scheduling of Public Transport 2*. North-Holland
19. Lessard, R., Rousseau, J.M. & Dupuis, D. (1981). HASTUS I: A mathematical programming approach to the bus driver scheduling problem. In A. Wren (Ed.) *Computer Scheduling of Public Transport: Urban Passenger Vehicle and Crew Scheduling*. North-Holland, pp. 255-268
20. Mason, A.J., Ryan, D.M. & Panton, D.M. (1998). Integrated simulation, heuristic and optimization approaches to staff scheduling. *Operations Research*, Vol. 46, pp. 161-175
21. Patrikalakis, I. & Xerocostas, D. (1992). A new decomposition scheme of the urban public transit scheduling problem. In M. Desrochers & J.M. Rousseau (Eds.) *Computer-Aided Transit Scheduling*, Lecture Notes in Economics and Mathematical Systems 386. Springer-Verlag, pp.407-425
22. Ryan, D.M. & Falkner, J.C. (1987). On the integer properties of scheduling set partitioning models. *European Journal of Operational Research*, Vol. 35, PP. 442-456
23. Ryan, D.M. & Foster, B.A. (1981). An integer programming approach to scheduling. In A. Wren (Ed.) *Computer Scheduling of Public Transport: Urban Passenger Vehicle and Crew Scheduling*. North-Holland, pp. 269-280
24. Schrage, L. (1999). *Optimization Modeling with LINGO*, 3/e. Lindo Systems Inc.
25. Segal, M. (1974). The operator-scheduling problem: a network flow approach. *Operations Research*, Vol. 22, 808-823
26. Valouxis, C. & Housos, E. (2002). Combined bus and driver scheduling. *Computers & Operations Research*, Vol. 29, 243-259
27. Vance, P.H., Barnhart, C., Johnson, E.L. & Nemhauser, G.L. (1997). Airline crew scheduling: a new formulation and decomposition algorithm. *Operations Research*, Vol. 45, pp. 188-200
28. Wren, A., Smith, B.M. & Miller, A.J. (1985). Complementary approaches to crew scheduling. In J.M. Rousseau (Ed.) *Computer Scheduling of Public Transport 2*. North-Holland, pp. 263-278

An Interactive Approach to Fuzzy-based Robust Design

Hideo Fujimoto, Yu Tao, Satoko Yamakawa

Department of System Engineering, Nagoya Institute of Technology,
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
E-mail: fujimoto, tao@vier.mech.nitech.ac.jp

Abstract

In order to win in the violent global competition, it is necessary for vendors to shorten the product development of novel, better product on quality consideration. It is expected to be an efficient way to extend robust design to the early stage of design. In this paper, we define multiple aspects of robust design from the view of multi-objective optimization. For the existing two problems in related researches, a new robust design model is proposed and embedded into an interactive approach as a decision making procedure. Fuzzy LR Number is used in this model to deal with asymmetric distributed variables and capture the imprecision in design variables in the early phases of design. Assisting by the proposed approach in which a proposed reference mechanism is embedded, designer is available to use progressively generating preference structure to select satisfactory solution from the optimized Pareto solution set. Pressure vessel design is used as an example to demonstrate the proposed approach.

Key Words: Robust Design; Interactive Approach; Fuzzy LR Number; Decision Making; Pareto Solution

1. Introduction

Every engineering design is inevitably subject to variability that arises from a variety of sources, such as manufacturing process, it is necessary to firstly achieve the robustness of design in the preliminary stage of design process. Disregarding uncertainty will cause rejected parts, high manufacturing cost, or failure in use and service. Elimination of uncertainty in the above sources is usually expensive or impossible in technique aspect. Robust design is regarded as an efficient tool to evaluate uncertainty and find relatively robust solution to existing uncertainty.

Taguchi's robust design methods have been widely used to improve design quality of products and processes. From the concept of Taguchi's method, the quality of a product is improved by minimizing the effect of the cause of variation without eliminating the causes[1,2]. Although the methods Taguchi offered for robust design have received some criticisms[3], fundamental principles and concepts of robust design have been widely accepted. One of them is that variation of performance decides loss in quality and tradeoff between performance and variation of performance will determine the quality of design. This philosophy can be addressed as a multiobjective optimization problem.

There are two problems existed in previous researches. One is how to properly quantify uncertainty in design variables and parameters. However, normal distribution widely used in most of research is not adequate[4-7]. Another problem is more difficult and challenging -- how to clearly express the designer's preference structure for optimal solution.

We use Fuzzy LR number to solve the first problem. Though Arakawa introduced Fuzzy theory into robust design [8] to express the imprecision in robust design, how to get the robust optimal from Pareto solutions has not been adequately addressed in their research, and separately maximizing variation of design variable causes critical technical obstacle in performing optimization.

To solve the second problem, a feasible way is to create an Pareto alternative section for selection, and designer select the "relatively best" one from it. Designer's preference information is generated progressively based on optimized results from the previous step. This is an interactive approach.

2. Proposed Approach

2.1 Multiple Aspects of Robust Design

Taguchi uses the quality loss function as a metric for robust optimization. It is from the view of probabilistic[1]. Using the Taguchi's robust design conception, it is necessary to model multiple aspects of robust design fully as at least two objective functions. Arakawa addressed the features of robust design as three rules and formulated from the view of Fuzzy theory.

In this paper, we consider the robust design problem as the following four aspects:

1. solution should be the satisfactory one selected from the Pareto optimal set;
2. the closer is the solution to the utopia point, the better;
3. the expected performance and the variation of performance should be minimized simultaneously;
4. variables should be selected with considering flexibility in manufacturing.

Based on these, we proposed to deal with robust design problem as a Tri-Objective Decision Making (TODM) problem.

2.2 Proposed Robust Design Model

From the above four aspects, we formulate robust design as the following TODM problem:

$$
\begin{aligned}
&\text{Find } \tilde{x} = \left(x_i, x_i^L, x_i^R\right) \subset D \qquad ,(i = 1,2,\cdots,n) \\
&\text{for given } \tilde{z} = \left(z_k, z_k^L, z_k^R\right) \qquad ,(k = 1,2,\cdots,m) \\
&\text{minimize } \frac{f(\tilde{x},\tilde{z})}{f^*(\tilde{x},\tilde{z})} \\
&\text{minimize } \frac{\Delta(\tilde{x},\tilde{z})}{\Delta^*(\tilde{x},\tilde{z})} \\
&\text{maximize } \sum_{i=1}^{n} \frac{(x_i^L + x_i^R)}{x_i^*} \\
&\text{subject to} \\
&\qquad g_j(\tilde{x},\tilde{z}) + \Delta g_j(\tilde{x},\tilde{z}) - g_{ja}(\tilde{x},\tilde{z}) - \Delta g_{ja}(\tilde{x},\tilde{z}) \le 0 \\
&\qquad \Delta g_j(\tilde{x},\tilde{z}) = \sum_{\frac{\partial g}{\partial x_i} > 0} \frac{\partial g_m}{\partial x_i} \cdot x_i^R - \sum_{\frac{\partial g_m}{\partial x_i} < 0} \frac{\partial g_m}{\partial x_i} \cdot x_i^L \\
&\qquad\qquad + \sum_{\frac{\partial g_m}{\partial z_k^R} > 0} \frac{\partial g_m}{\partial z_k^R} \cdot z^R + \sum_{\frac{\partial g_m}{\partial z_k^L} < 0} \frac{\partial g_m}{\partial z_k^L} \cdot z^L \\
&\qquad x_i^L + (\Delta x_i^L)^U \le x_i \le x_i^R - (\Delta x_i^U)^L
\end{aligned}
\qquad (1)
$$

where $\tilde{x}$ and $\tilde{z}$ are design variables and uncontrollable parameters respectively. x_i^L, x_i^R are the lower and upper bound of x_i which expressed using triangle Fuzzy LR Number. $f^*(\tilde{x},\tilde{z})$, $\Delta^*(\tilde{x},\tilde{z})$ are the individually optimized solutions of taking performance $f(x,z)$ and variation of performance $\Delta(x,z)$ as objective function respectively. D called the feasible region which means design variable satisfy all of constraints. x_i^* is a normalized factor which uses the optimal solution without considering the uncertainty in objective functions and constraints. $\Delta g_{ja}(\tilde{x},\tilde{z})$ is the upper bound of allowable limit of the j-th constraint $g_{ja}(\tilde{x},\tilde{z})$. $(\Delta x_i^U)^L, (\Delta x_i^L)^U$ are the Fuzzy left and right number, which represent lower and upper limit of variation Δx_i^U, Δx_i^L. Constraints are formulated in the worst case which assumed all variations of system performance will occur simultaneously in the worst possible combination of design variables. [$f^*(\tilde{x},\tilde{z})$, $\Delta^*(\tilde{x},\tilde{z})$] can be denoted as utopia pint of Pareto solution set.

A robust design problem is modeled as a TODM problem, in which performance and variation in performance are minimized to obtain robustness of performance, while total sum of variations of variables are maximized for relaxation in acceptable range of variations in manufacturing.

Using the proposed model, decision making is proposed to be embodied in an interactive process.

2.3 ε - constraint interactive procedure

In this paper, the interactive approach developed by Palli[9] is improved.
The ε-constraint method can be formulated as

$$\begin{aligned} &\text{minimize} \quad f_1(\mathrm{x}) \\ &\text{subject to} \quad f_i(\mathrm{x}) \le \varepsilon_i, i = 2, \cdots, m, \\ &\qquad g_j(x) \le 0, j = 1, \cdots, k \\ &\qquad \mathrm{x} \in D \end{aligned} \tag{2}$$

where the most principal function $f_1(x)$ is still taken as primary objective function, while other objective functions are taken as constraints.

When the upper limit ε_i of i-th converted objective function is decided, the value of primary function will be available to be identified. So even spread Pareto solutions can be achieved even in concave case.

2.4 Robust Design Procedure through an Interactive Approach

The proposed procedure employs an interactive procedure in which the designer's preferences are obtained progressively.

The interactive approach which embedded robust design consists of six steps:

Step1-- problem formulation

To formulate a certain robust design problem as a TODM problem as in Eq. (1) and consider it under the constraints from the worst scenario.

Step 2-- choose optimized candidate size P and a primary objective function

P means the number of optimized candidates presented to designer to make tradeoff in every stage.

Step 3-- determine limits of the objective functions

Designer has to determine the upper and lower limits of objective functions in the first stage. Since the objective functions are normalized, the lower limits are greater than 1. Or designer can simply use the utopia point as the lower value of design variables. But the initial upper limit has to be determined vaguely according to actual technical context.

Step 4-- convert to a constrained form

In order to ensure Pareto solutions are widely dispersed, the clustering technique [10]can be used here.

For i=2,...,m,

$$\varepsilon_i = \varepsilon_{i,lower} + \alpha\left(\varepsilon_{i,upper} - \varepsilon_{i,lower}\right) \tag{3}$$

where α is a random number between 0 and 1.

Then the secondary objective functions can be converted into the constraint form. This step is repeated P times to obtain P Pareto solutions.

Step 5-- present solutions to the designer

The obtained Pareto solutions are provided to designer through a graphical user interface. The most preferred solution may be the best value of the primary objective function while the values of the other objective functions are acceptable.

We propose a reference choice mechanism, which use the information of scenarios in current stage. One scenario with the least sum of squares of deviation is recommended as the reference one.

$$ave = \frac{1}{Pm}\sum_{i=1}^{P6}\sum_{j=1}^{m} obj(i,j)$$

$$sqdev(i) = \sum_{j=1}^{m}(obj(i,j) - ave)^2, i = 1,\cdots,P \quad (4)$$

where *ave* is the mean of m objective functions in current stage, *sqdev(i)* is sum of squares of deviation of m objective functions in i-th scenario. With reference mechanism, designer can perform interactive approach efficiently.

Once the preferred scenario is chosen, the new ε-interval is generated around the chosen point. A reduction factor $r(r>1)$ is chosen to decrease the each ε-interval. Steuer gave a recommended value for r :

$$(1/p)^m \le r \le \nu^{1/(H-1)} \quad (5)$$

where, ν is the final length ,and H is the number of total interactive times which designer prefers.

The generation of new ε-interval was discussed by Palli[9] as the following :

$$\Delta\varepsilon_i = \varepsilon_{i,upper} - \varepsilon_{i,lower}, \varepsilon_{i,upper} = \varepsilon_i + \frac{\Delta\varepsilon_i}{r}, \varepsilon_{i,lower} = \varepsilon_i - \frac{\Delta\varepsilon_i}{r} \quad (6)$$

where $\Delta\varepsilon_i$ is the length of the original interval for $f_i(i = 2,\ldots,m)$. The new ε_i interval is generated around the current ε_i.

Step 6-- check for convergence

If the solution is converged to an acceptable preferred solution or ε-interval is less than some prescribed small value, the iteration process stops. Otherwise, steps 3-5 are repeated and another set of P Pareto optimal points are generated and presented to the designer.

If the solutions obtained are not satisfactory to the designer at the beginning of step 5, the clustering procedure in step 4 can be repeated to obtain a new set of Pareto solutions for the objective functions.

3. Pressure Vessel Design Problem

We use a simple pressure vessel design problem [11] to demonstrate the advantage of the proposed procedure. Motivation is that the total manufacture cost is to be minimized. Design variables are : x_1 and x_2, the thickness and length of shell, x_3 and x_4, the thickness and diameter D of the head.

The problem is formulated as follows:

$$f(x;p) = 2c_w\pi\left(\frac{x_1}{\sin p_2}\right)^2\left(\frac{p_2}{360}\right)p_1x_3 + 4c_w\pi^2\left(\frac{x_1}{\sin p_2}\right)^2\left(\frac{p_2}{360}\right)p_1x_2 + 2c_s\pi p_1x_1x_2x_3 + 2c_h\pi p_1x_4x_2^2 \tag{7}$$

where c are cost factors[11].

Assumed the case without sufficient statistical information about design variables and parameters, we use Fuzzy LR number to express the evaluation of variation of design variables and tabulated in Table 1. Asymmetric distribution of variation of variables is considered according to actual needs from manufacturing.

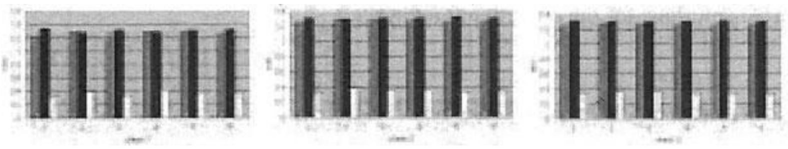
Table 1. initial nominal value and range of variables

	M	LB	UB	LBL	UBL	LBR	UBR
x_1	3.0	0.0	0.09	2.7e-3	4.5e-3	2.7e-3	4.5e-3
x_2	305	9.15	15.25	0.183	0.366	0.183	0.366
x_3	1.5	0.0	0.045	9e-1	1.5e-3	9e-4	1.5e-3
x_4	125	3.75	3.75	0.188	0.188	0.188	0.188

In this example, parameters p_1 and p_2 are density and welding angle respectively. Variation of parameters are assumed as within 5% of the corresponding mean value.

Objective function of total sum of variations of variables is chosen as the primary function, and P is determined as 6 in every stage. Objective functions of cost and variation of cost are converted into ε-interval constraints. In order to decide the lower and upper bound of converted constraints, individual optimization is performed for objective functions of cost and variation in cost with consideration of uncertainty in constraints. Optimization is performed to decide ideal variables x_i^* without consideration of uncertainty in objective functions and constraints. Ideal variables ={2.8575,298.958,1.5875,121.158}, and lower bound of cost and variation in cost are 7198.20 and 294.51 dollars respectively. Upper limits are assumed as 5 times of corresponding lower limits.

Then optimized results are presented to designer for making tradeoff as in Fig.1.



(a) initial stage **(b) second stage** **(c) last stage**

Fig. 1. stages in proposed interactive approach

Blue stick represents cost, red one variation of cost, and yellow one global variation of variables in Fig.1.

Based on reference mechanism described in Eq. (4), the referred solution is the 5th, 3rd, and 2nd one in sequence from 1st stage to last stage. However, its potential assumption is the same priority among objective functions. The referred solution can be taken as a benchmark at least. Designer can choose satisfactory solution according to his preference. And around the chosen solution, Pareto candidate section will be extended in the next stage. Till the 3rd stage, all candidates are nearly same and is a hint of convergence. For its random nature of approach, it is reasonable to take this convergence point as a global optimal one.

Besides the optimized value of objective function, some useful information can be found in the optimized variables and variation of variables. Applying general robust design model [2,7], both of the lower limits and upper limits tend to the prescribed minimum limit. Compared with general robust design model [2,7], optimized upper limits of all variables are greater than those of other models. Compared with optimized mean value, high percentage of upper limit of optimized variation of variable means cheap material may be used or relax in manufacturing and check. Low percentage means stern management of quality has to be performed. Of course, high percentage is desired. In this example, x_4 is at low percentage (2%)and needed stern quality management, while x_2 is at high percentage (6.4%) and will tolerate more random variation in length dimension.

4. Conclusions

In this paper, the proposed robust design model is embodied in an interactive procedure to help designer in finding satisfying solutions from Pareto solutions.

There are several major advantages of using proposed approach:

- To assist designer in selecting satisfied solution from candidates of Pareto solution set using even originally ambiguous preference through progressively interactive approach. Reference mechanism is proposed to provide designer a benchmark to select efficiently.
- Asymmetric distributed variables can be dealt with using Fuzzy LR number. Instead of normal distribution, Fuzzy LR number shows designer more useful information, such as the sensitive tendency of tolerance of variables. It is helpful for designer to escape potential sensitive risk and obtain robust solution.
- Proposed approach provides flexibility and guideline for manufacturing by maximizing total sum of variations of variables. It is helpful in extending the acceptable tolerance range of variables which is too narrow optimized by general robust design models. It is also helpful in reducing the potential risk to critical accumulated tolerance and harmful collision in service caused by over-relaxed tolerance range.

In this paper, for illustration purpose, a simple pressure vessel design problem is introduced. From the results of example, the validity of the proposed method has been demonstrated.

References

[1] Taguchi, G., 1991, From static Type S/N Ratios to Dynamics Type S/N Ratios, *Standardization Qual. Control*, 44(9),52-59.

[2] Parkinson, A., Sorensen, C., and Pourhassan, N., 1993, "A General Approach for Robust Optimal Design," *Transactions of the ASME*, Vol. 115, pp. 74-80.

[3] Antonsson, E.K., and Otto, K.N., 1995, "Imprecision in Engineering Design," *Transactions of the ASME*, Vol. 117, pp. 25-32.

[4] Ramakrishnan, B., and Rao, S.S., 1991, "A Robust Optimal Approach Using Taguchi's Loss Function for Solving Nonlinear Optimization Problems," *Advances in Design Automation*, ASME DE-32-1, pp. 241-248.

[5] Bras, B.A., and Mistree, F., 1995, "A Compromise Decision Support Problem for Robust and Axiomatic Design," *ASME Journal of Mechanical Design*, Vol. 117, No. 1, pp. 10-19.

[6] Chen, w., Allen, J.K., Mistree, F., and Tsui, K,-L., 1996b, "A procedure for Robust Design: Minimizing Variations Caused by Noise Factors and Control Factors," *ASME Journal of Mechanical Design*, Vol.118, pp. 478-485.

[7] Chen, w., Allen, J.K., and Mistree, F.,1997, "The Robust Concept Exploration Method for Enhancing Concurrent Systems Design," Journal of *Concurrent Engineering: Research and Applications*, Vol.5, No. 3, pp. 203-217.

[8] Arakawa, M., Yamakawa, H., and Hagiwara, I., 1999, "Consideration of Robust Design Methodology by Using Fuzzy Numbers," *Transactions of the JMSE*, Vol. 65-632C, pp.297-304.

[9] Palli,N.,et.al, 1998, "An Interactive Multistage ε-Inequality Constraint Method for Multiple Objectives Decision making," *Transactions of the ASME,* Vol. 120, pp.678-686.

[10] Miettinen, K., 1997, "Review of Nonlinear MCDM methods," 6th International Summer School of MCDA, Turku, Finland.

[11] Kannan, B. K., Kramer, S. N., 1994, "An Augmented Lagrange Multiplier Based Method for Mixed Integer Discrete Continous Optimization and Its Applications to Mechanical Design," *ASME Journal of Mechanical Design*, Vol.116, pp. 405-411.

A Hybrid Genetic Algorithm for Solving a Capacity Constrained Truckload Transportation with Crashed Customer

Sangheon Han[1] and Yoshio Tabata[2]

Graduate School of Economics, Osaka University, Machikaneyama 1-7, Toyonaka, Osaka, JAPAN

Abstract. Disaster has often led to many damage points throughout total transportation network. In this paper we deal with the capacity constrained truckload transportation problem with a post-crashed point. Our objective is to reconstruct an optimal route to minimize the number of truck and vehicle movements after occurring a post-crashed point. We propose a Hybrid-Genetic Algorithm consisting of ordinary GA and Sweeping algorithm. A detailed numerical study is conducted and its results show the advantages of our proposed algorithm comparing with other heuristic. Moreover, we demonstrate the optimal GA-parameters setting using the design of experiments for efficiency of our algorithm.

1 Introduction

As for the transportation network in the world, many disaster (earthquake, storm, accident and others) may always happen, and hampering the transit for basic life and economical activities. In this paper, we consider the case crashed customer model in the Vehicle Routing Problem (VRP). To promote reconstructing effectiveness, well considered reconstruction plan for the road-network has to be considered.

First of all, the damage characteristics of road-network are taken into consideration. The reconstruct task for each damage point (crushed customers) should be reasonably assigned so as to maximize the reconstruction effectiveness. The effectiveness of reconstruction can be maximized by minimizing the total travelled distance or time.

Secondly, the schedule model for reconstruction is a combinational optimization problem. In actually, the schedule model is the expansion of work assignment model to incorporate the desired object, e.g., minimizing the total working time[6,9]. Relative researches by many authors could be easily found, for example, the scheduling problem solved by branch and bound method [3, 5], by genetic algorithm approach [1, 2, 4], etc..

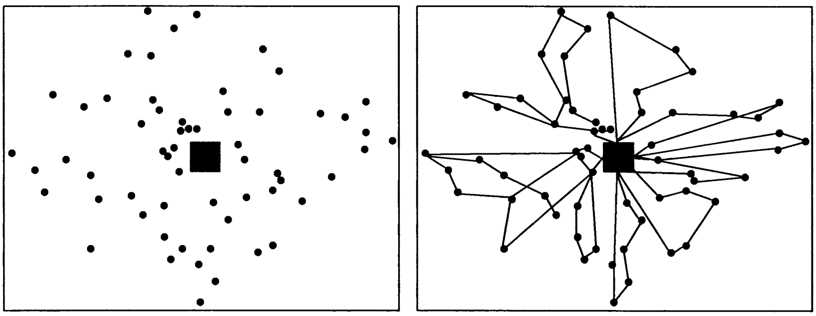
Thirdly, this study is far differnt from aforementioned traditional models in some aspect: Our problem is the case that customers get damaged by disasters; thus, our problem has more complexity than others. For example, if one customer is crashed by disaster, liked all routes will be disappeared. Morever, the VRP is well-known NP-hard problem. In view of the aforementioned reasolution complexity, the concept of genetic algorithm (GA) is applied in this study as to obtain a heuristic solution [6, 16]. Although our problem is a well-known combinatorial optimization problem, but a satisfying heuristic solution can be more easily and efficiently derived by the recent developted GA [15, 25, 26].

Fianlly, the A simple road network will be illustrated as a numerial example to validate our model. Study show that a satisfying solution can be efficiently derived by our modified Hybrid Genetic Algorithm (HGA). Thus, this study can be a powerful basis for the case of crashed customers simulation.

2 The Vehicle Routing Problem; the case of crashed customers

2.1 Classical VRP

The Vehicle Routing Problem (VRP) was originally proposed by Dantzig and Ramser [5] and defined as follows: vehicles with a fixed capacity Q must deliver order quantities q_i $(i = 1, ..., n)$ of goods from a single depot $(i = 0)$ to n customers. Knowing the distance d_{ij} between customers i and j $(i, j = 0, ..., n)$, the objective of the problem is to minimize the total distance travelled by the vehicles in such a way that only one vehicle handles the deliveries for a given customer and the total quantity of goods that a single vehicle delivers is not larger than Q.



(a) A Vehicle Routing Problem (b) A possible solution

Fig. 1. The example of VRP

Fig. 1. gives a graphical representation of a VRP and one possible solution. The square (in the middle of Fig 1(a) and (b)) represents the base (where the trucks start and finish their tour) and the diamonds represent the sub-routes.

Figure 1(b) shows the tours of the different trucks. It should be observed that, in this case, all the customers have been allocated.

Problem formulation

Let $G = (V, A)$ be a graph with a set V of vertices and a set A of arcs. We have $V = 0 \cup N$, where 0 corresponds to the depot and $N = 1, ..., n$ is the set of customers. For the set of arcs, we have $A = (\{0\} \times N) \cup I \cup (N \times \{0\})$, where $I \subseteq N \times N$ is the set of arcs connecting the customers, $\{0\} \times N$ contains the arcs from the depot to the customers, and $N \times \{0\}$ contains the arcs from the customers to the depot. Every customer $i \in N$ has a positive demand q_i. For each arc $(i, j) \in A$ we have a cost c_{ij}. Furthermore, we assume that the vehicles are identical and have the capacity Q. All the above mentioned factors are assumed to be known in advance. Thus the model examined is deterministic.

We have the following variables: For each customer $i \in N$, y_i is the load of the vehicle when it arrives at the customer. Now the problem is to determine which of the arcs $(i, j) \in A$ are used by routes. For each arc $(i, j) \in A$, the decision variable x_{ij} is equal to 1 if arc (i, j) is used by a vehicle and 0 otherwise. Formally

$$\text{Minimize} \quad \sum_{(i,j)\in A} c_{ij}x_{ij} \tag{1}$$

$$\text{subject to} \quad \sum_{j\in V} x_{ij} = 1 \qquad \forall i \in N \tag{2}$$

$$\sum_{j\in V} x_{ji} = 1 \qquad \forall i \in N \tag{3}$$

$$x_{ij} = 1 \Rightarrow y_i - q_i = y_j \qquad \forall (i, j) \in I \tag{4}$$

$$q_i \leq y_i \leq Q \qquad \forall i \in V \tag{5}$$

$$x_{ij} \in \{0, 1\} \qquad \forall (i, j) \in A \tag{6}$$

We minimize the total costs that consist of travel costs and a fixed cost c of vehicles (included in the travel cost c_0 between depot and first customer). The object is, firstly minimize the number of routes or vehicles, and then the total distance of all routes. By equation (2), (3) and (6), we require that every customer be visited exactly once. Equation (4), (5) enforce that the loads of the vehicles when arriving at the customers are feasible.

2.2 VRP with crashed customers

Basic model structure is similar to classical VRP model. However, this crashed customer problem is the special case in the VRP. According to the crash of customers in the road-network, the situation varying with time can be considered. Meanwhile, this crashed customer case can be thought the case that customer canceled his order for goods. The example illustrated as below.

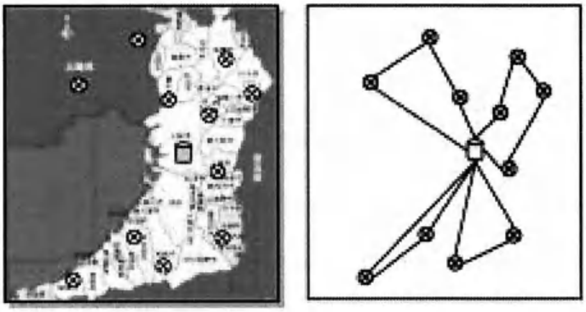


Fig. 2. Distribution Network in Osaka, Japan

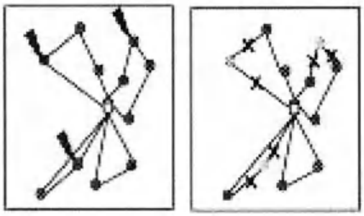


Fig. 3. crashed points and pathes

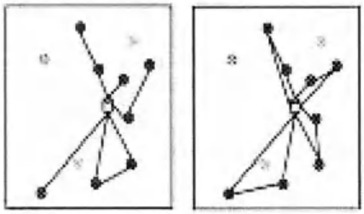


Fig. 4. reconstruct new path

In upper example, Fig.2. is the map of the distribution network in Osaka, Japan. The shape of cylinder indicate depot and circle indicate customers. In Fig.3., suppose that three customers crashed and crashed simultaneously with liked all routes. Reconstruction in the road-network will be dynamically updated as the Fig.4.. When we reconstruct new route after crashed, we must consider the cost (in this study, the cost is traveled distance). So, we re-calculate the new-route, considering the present routes (Method-2). That is, calculate the new route from the rest routes. The another method is full calculation from the first (Method-1). We will compare these two method in numerical examples on the section 4.

3 A hybrid methodology for Vehicle Routing Problem

The basic concept behind the hybrid methodology is not to use the GA to directly optimize the parameters of the solution, but rather to use the GA to optimize the parameters of a simple heuristic problem solving strategy. The approach is shown in the following diagram Fig.5. The representation of a feasible solution in a chromosome structure may be much more complex for the VRP than other problems. In addition to the problem of finding an optimal route for each vehicle, there is also the problem of distributing the number of visits required by each customer in the planning horizon, while

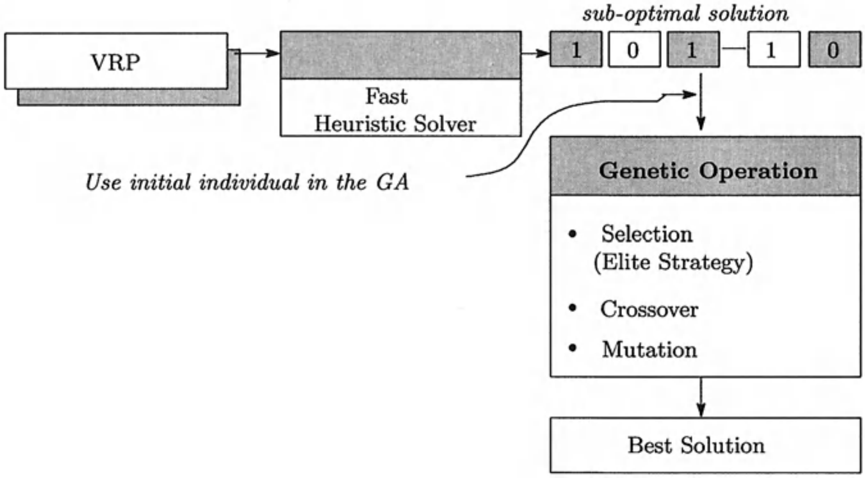


Fig. 5. Hybrid Genetic Algorithm for VRP

satisfying all the constraints. Meanwhile, GA operator (crossover, mutation) make lethal genes in VRP generally. To control lethal gene, we developed a HGA with controlling lethal gene[].

(1) Gene type

If we have the following solution:

Route No. 1 is $0 \to 1 \to 2 \to 0$
Route No. 2 is $0 \to 3 \to 4 \to 5 \to 0$
Route No. 3 is $0 \to 6 \to 7 \to 0$

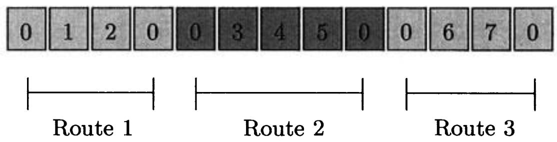


Fig. 6. Chromosome representation

(2) Crossover

To prevent invalid offsprings (lethal genes) from being reproduced, we propose an order based crossover operators as below.

Let P_1 and P_2 be the parent strings, $P_1[1], ..., P_1[n]$ and $P_2[1], ..., P_2[n]$, respectively. And, let C be the child string. Assuming a minimization problem, then for all $i = 1, ..., n$:

S0 Let $U[\cdot] = [1, \ldots, i, i+1, \ldots, n]$ as set of all customers.
S1 Select any customer i in U.
S2 Pick out subroutes including i from P_1 and P_2, respectively.

$S_1[\cdot]$: subroute of P_1
$S_2[\cdot]$: subroute of P_2

S2-1 If $S_1[\cdot]$ and $S_2[\cdot] \subset U$,
then the one is selected of them with below probabilities.
Probability to select $S_1[\cdot] = f_2/(f_1 + f_2)$
Probability to select $S_2[\cdot] = f_1/(f_1 + f_2)$
f_1 and f_2 are the number of times that $S_1[\cdot]$, $S_2[\cdot]$ are selected as subroute of C, respectively.

S2-2 If $S_1[\cdot] \subset U$ or $S_2[\cdot] \subset U$,
then the subset inclued in U is used for C as subset.

S2-3 If $S_1[\cdot] \not\subset U$ and $S_2[\cdot] \not\subset U$,
then both $S_1[\cdot]$ and $S_2[\cdot]$ are not selected.

S2-3-1 Select any customer j in U at random.
Let $S[\cdot]$ is subroute of C.

S2-3-2 $S[\cdot] + [j] \rightarrow$ Calculate amount of deliveries
If amount of deliveries of $S > Q$ (vehicle capacity),
then end.
else $S[\cdot] + [j]$ then return to **S2-3-1**.

S2-3-3 If U becomes a null set, then end.

S3 Omit $S[\cdot]$ (decided upper procedure) in U and then, go to **S1**
If U becomes a null set, then end.

To generate another offspring, the upper procedure performs one more time.

(3) Mutation

The mutation operator adopted in this paper is as follows :

S0 Select any two customers i, j in the $P[\cdot]$ at random.

S1 Try swap i for j

$\downarrow$

If subroutes including i or $j < Q$

$\downarrow$

then swap i from j
else not swapping

end

4 Numerical Example and Discussions

As a assumed VRP with crashed customer in Fig. 2. is used to validate our HGA. The reconstruct effectiveness for a damage customer depends on the number of defeated customer. The results of simulation is shown in Table.1..

From result in the table 1., over almost all the problem, Method-2 shows better than method-1 almost. Problem 3 and problem 5 are not better than initial results. Maybe, it is because of the random property of GA and the seisitivity of GA-parameters. However, we can control the parameters by the design of experiments method[].

Table 1. Results for various customer sizes

	Before Crash	Post-crashed Customers	
	Initial Result	Method 1	Method 2
Problem 1	1344.22	1681.984	1699.097
Problem 2	590.2655	651.8612	713.1203
Problem 3	460.9023	458.8078	514.5714
Problem 4	332.513	332.918	395.0115
Problem 5	249.6133	243.127	284.8044
Problem 6	199.9878	196.6487	237.6387

All results are the value of fitness.

5 Conclusions and Recommendations

VRP is one of the classic problems associated with TSP, and it has been applied in various fields. Since VRP is difficult to solve in real time, heuristic methods have been adapted to solve it. We developed an efficient Sweep-Genetic Algorithm for the special type of VRP and then, we could get good results from the proposed method. Morever, the design of experiments method makes we identify the correlation of GA-parameters (crossover probability and mutation probability) and decide adaptable probabilities. Furthermore, we compared method-1 with method-2 mentioned on section 2.2 for our problem using Hybrid-GA. Since GA has Randomize property, it cannot guarantee the result in general, but the proposed method HGA could find the optimal solution or at least sub-optimal. GA and Hybrid-GA are valid realistic VRP as well as the homogeneous VRP. We believe that our method can easily be adapted to solve the realistic problem.

References

1. C.R. Houck, J.A. joines, and M.G. Kay, "Comparison of genetic algorithms, random restart and two-opt switching for solving large location-allocation problems", Computers and Operational Research, vol.23, no.6, pp.587-596, 1996.
2. D.G. Gonway and M.A. Venkataramanan, "Genetic search and the dynamic facility layout problem", *Computers and Operations Research*, vol.21, no.8, pp.955-960, 1994.
3. E.Demeulemeester and W. Herroelen, "A branch and Bound procedure for the multiple resource-constrained project scheduling problem", *Management Science*, vol.38, no.12, pp.1903-1818, 1992.
4. G. Dantzig and J.H. Ramser 1959. The Truck Dispatching Problem, *Management Science* 6(1), pp. 80-91.

5. H.Ishibushi, T.Yamamoto, and H. Tanaka, "Genetic algorithms and neighberhood search algorithms for fuzzy flow-shop scheduling problems", *Fuzzy Sets and Systems*, vol.67, no.3, pp81-100, 1994.
6. J.P. Stinson, "Multiple resource-constraint scheduling using branch and bound", *AIIE Transaction*, vol.10, no.3, pp.252-259, 1979.
7. J. Weglarz, "Control in resource allocation systems", *Fonundation od Control Engineering*, vol.5, no.3, pp.159-180, 1980.
8. M. Sakawa, M.Inuiquchi, H. Sunada, and K. Sawada, "Fuzzy Multiobjective Combinatorial optimization through Revised Genetic Algorithm", *Journal of japan Society for Fuzzy Theory and Systems*, vol.6, no.1, pp.177-185, 1997.
9. M. Sakawa, K. kato, H. Sunda, and T. Shibano, "Fuzzy Programming Problems through Revised Genetic Algorithms", *European Journal of Operational Research*, vol. 97, no2, pp.149-158, 1997.
10. O. Linet and U. Gunduz, "A survey on resource-constrained project scheduling problem", *IIE Transactions*, vol.29, no.5, pp.574-586, 1995.
11. S.HAN 2002 Design of experiments to identify optimal parameters of genetic algorithm. *Journal of Scientiae Methematica Japonica*, 55,No.3(2002), 539-545, :**e5**, 281-287.
12. Z. Michalewicz, *Genetic algorithms + Data Structures = Evolution Programs*, Springer Verlag Press, Berlin, 1996.

A Multi-Objective Programming Approach for Evaluating Agri-Environmental Policy

Kiyotada Hayashi

National Agricultural Research Center, Tsukuba 305-8666, Japan

Abstract. A multi-objective programming approach is presented for evaluating agri-environmental policy, in which various management practices are introduced to reduce environmental problems caused by, for example, fertilizer and pesticide application. Particular attention is paid to environmental impacts of agricultural practices. After comparing with an economic approach based on mathematical programming, possibility of integrating various kinds of indicators is discussed. Some implications of constructing integrated evaluation models are also considered.

1 Introduction

Various agricultural systems have been analyzed by mathematical programming. Cropping, fertilizer application, pest and disease control, livestock feeding, and livestock breeding and replacement have been modeled as mathematical programs [3]. In addition, planning models in which risks are incorporated are commonly used in agricultural economics and farm management [4] as well as in, for example, finance. The pervasiveness of the programming is reflected in the fact that linear programming is taught as an indispensable tool for farm management [8]. One of the recent trends in mathematical programming applied to agricultural systems is that the number of applications of multi-objective programming (including goal programming) is increasing [5].

As the increase in number suggests, considering multiplicity of evaluation criteria has become important in analyzing agri-environmental issues. For example, as a result of the seriousness of environmental degradation caused by intensive farming practices, the trade-offs between agriculture and the environment have attracted public attention. A typical example of the environmental problem is water pollution (the nitrate issue) caused by chemical fertilizers and manure. Moreover, introduction of agri-environmental programs in, for example, Western Europe has increased public interest in the evaluation of the program and thus necessitated the discussion on how to construct a variety of criteria for the evaluation. There are, however, difficulties in coping with multiple objectives in the agri-environmental problems, because the problems are related to a wide range of phenomena and the problems have complicated cause-effect relationships.

In this paper, therefore, we review the appropriateness of a multi-objective programming approach to the evaluation of agri-environmental policy by comparing with an economic approach based on mathematical programming and by constructing a conceptual model for environmental impacts of agricultural practices.

2 Mathematical Programming Approach to Agri-Environmental Problems

In this section, the main feature of a multi-objective programming approach is outlined through the comparison with an economic usage of mathematical programming in the area of agricultural economics. Then some issues we face in modeling agri-environmental problems are presented.

The mathematical programming approach based on micro economics is positive mathematical programming (PMP), which is developed by Howitt [7]. The model has a non-linear (quadratic) function instead of a linear objective function and has a capacity to overcome the calibration problem, the difficulty of programming models in calibrating against a base year or an average over several years. This approach has been applied to the evaluation of agri-environmental policies [12].

As contrasted with the approach, the fundamental characteristic of a multi-objective approach to agri-environmental problems is explicit modeling of multiple objectives. That is, trade-offs between agricultural production and the environment are articulated in modeling. For example, Lakshminarayan et al. [9] applied a multi-objective approach to watershed-level cropping decisions. In addition to an economic objective (profit), soil loss, nitrate-N leaching, and Atrazine leaching, which were estimated by a physical simulation model (EPIC; Erosion Productivity Impact Calculator), were used as environmental objectives. Falconer and Hodge [1] used a planning model in which maximization of profit and minimization of a hazard score can be considered as objectives, although they did not explicitly formulate the model as a two-objective program. The impacts of pesticide taxation were analyzed by the model.

Because of the development of the methodology such as aspiration level approaches, which can cope with the situation where many objectives have to be included in the model, this direction of modeling will be useful for solving real-world problems. For example, such kind of methodology has been applied to land use planning using GIS [2], in which many objectives such as maximization of food output, maximization of net revenue, and minimization of erosion are included.

In agri-environmental problems, however, we have to confront the following issues. (1) Although maximum admissible concentrations of, for example, nitrate-N in drinking water are defined in regulations, it is still preferable to reduce the level of contamination because there is uncertainty about the impacts of nitrate on human health. (2) The trade-offs between economic returns and the levels of nitrate-N in groundwater may be difficult to understand for most decision makers as compared with, for example, the trade-offs between a salary and a vacation in a job decision.

Thus, in the next section an integration of agri-environmental indicators, which can be recognized as raw data in many cases, is illustrated as an alternative way of evaluation. In addition, appropriateness of criteria for evaluating agri-environmental issues is discussed with considering understandability of criteria.

3 Possibility of Integrated Evaluation

In this section, conceptual modeling is utilized to illustrate environmental impacts of agricultural practices. Conceptual models have been used in, for example, ecological risk assessments and serve the following three purposes [14,13]: (1) to help understand the situation being assessed and to make explicit assumptions concerning the situation; (2) to provide a communication tool between risk managers and stakeholders; (3) to provide a basis for organizing and conducting risk assessment.

Concept mapping is utilized in this study to clarify the following difficulties in constructing integrated evaluation models for environmental impacts of agricultural practices, although visual representation techniques such as flow charts have been used in conceptual modeling applied to ecological risk assessments. The first difficulty is related to the way of understanding the diverse environmental impacts of agricultural practices. As OECD [10] shows, environmental impacts of agriculture cover a wide range of aspects such as soil and water quality, biodiversity, and landscape. The second is how to recognize complicated cause-effect relationships. In assessing environmental impacts of agricultural practices long-term effects have to be taken into account in addition to short-term effects. Moreover, macro (global) effects have to be depicted in addition to micro (local) effects. For example, the emission of CO_2 from a specific area is a global problem from the outset, whereas nitrate concentrations in groundwater is a local problem that has global impacts as cumulative local problems. Concept mapping will be useful in considering these points.

Although the concept map constructed in the study [6] contains many agricultural practices, many indicators (intermediate variables), and risk concepts, only the impacts of fertilizer application are illustrated in Fig. 1. The impacts are related to the earth system such as climate change and ozone layer depletion, to ecosystems such as vegetation change and eutrophication, and to air and water quality. These impacts of agricultural practices are made through the nitrogen and carbon cycles in the atmosphere, soil, and water. The crucial point in this figure is that the impacts can be integrated into risk concepts. Ecological and human health risks are used in the modeling. Minimizing these risks can be considered as dual goals of laws and regulations that use risk assessment to inform decision making [11].

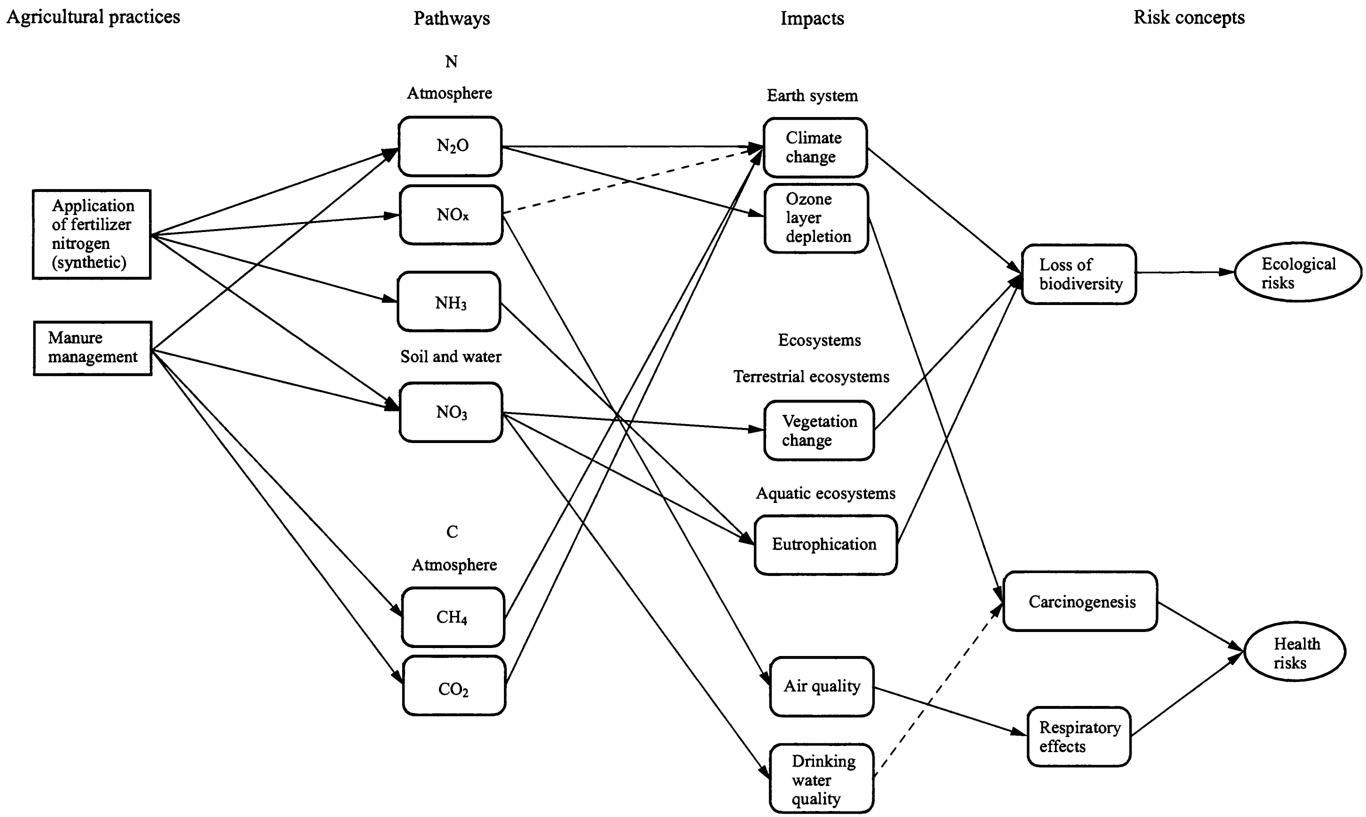


Fig. 1. An example of environmental impacts of agricultural practices

One of the important implications of the conceptual model is that it illustrates explicitly two kinds of modeling as shown in Table 1. That is, in Type A many indicators (intermediate variables) are used as evaluation criteria, whereas in Type B risk concepts are used and therefore a risk-benefit framework can be applicable.

Table 1. Examples of objectives used in two types of multi-objective models

Type A	Type B
(Max.) Economic return	(Max.) Economic return
(Min.) N_2O	(Min.) Health risks
(Min.) NOx	(Min.) Ecological risks
(Min.) NH_3	
(Min.) NO_3	
(Min.) CH_4	
(Min.) CO_2	

Since our primary concern is to make a conceptual model for environmental impacts of agricultural practices in order to think about the possibility of a multi-objective programming approach, although conceptual models can provide a basis for quantitative models, some theoretical considerations are given in the next section.

4 Concluding Remarks

A multi-objective programming approach based on Type B in Table 1 will help decision makers understand the meaning of trade-offs between evaluation criteria by relying on risk concepts using life-years such as disability adjusted life years (DALY) as well as on economic performance and thus it has the potential to be used in actual policy making processes, although numerical evaluation of risks may not be straightforward and expert knowledge may be necessary for estimating the numerical values. In interpreting properly the results obtained from the multi-objective approach, tripartition of methodology (normative, descriptive, and prescriptive) will be helpful, because the approach can be considered as a kind of decision analysis. This makes a striking contrast with the case of economic analysis, in which bipartition (normative and positive) is common. An implication of the difference in methodology for the evaluation of agri-environmental policy is that the meaning of calibration in the multi-objective approach becomes different from the case of calibration methods in economic analysis using mathematical programming because of the introduction of preference information into the model.

References

1. Falconer, K. and Hodge, I. (2001): Pesticide Taxation and Multi-Objective Policy-Making: Farm Modelling to Evaluate Profit/Environment Trade-Offs. Ecological Economics 36, 263–279
2. Fischer, G., Makowski, M. (2000): Land Use Planning. Wierzbicki, A. P., Makowski, M., Wessels, J. (Eds.): Model-Based Decision Support Methodology with Environmental Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands, 333–365
3. Glen, J. J. (1987): Mathematical Models in Farm Planning: A Survey. Operations Research 35, 641–666
4. Hardaker, J. B., Huirne, R. B. M., Anderson, J. R. (1997): Coping with Risk in Agriculture. CAB International, Wallingford, UK
5. Hayashi, K. (2000): Multicriteria Analysis for Agricultural Resource Management: A Critical Survey and Future Perspectives. European Journal of Operational Research 122, 486–500
6. Hayashi, K. (2002): Evaluating Farming Practices: Use of Health and Ecological Risk Concepts. Paper presented at the 13th Congress of the International Farm Management Association. Wageningen, The Netherlands. pp. 17 (`http://www.lei.dlo.nl/IFMA/files/papersandposters/PDF/Papers/HayashiK.pdf`)
7. Howitt, R. E. (1995): Positive Mathematical Programming. American Journal of Agricultural Economics 77, 329–342
8. Kay, R. D. (1994): Farm Management, 3rd Edition. McGraw-Hill, New York
9. Lakshminarayan, P. G., Johanson, S. R., Bouzaher, A. (1995): A Multi-Objective Approach to Integrating Agricultural Economic and Environmental Policies. Journal of Environmental Management 45, 365–378
10. OECD (2001): Environmental Indicators for Agriculture, Volume 3, Methods and Results. Paris
11. Presidential/Congressional Commission on Risk Assessment and Risk Management (1997): Risk Assessment and Risk Management in Regulatory Decision-Making. Washington, D.C.
12. Röhm, O., Dabbert, S. (1999): Modelling Regional Production and Income Effects. Van Huylenbroeck, G., Whitby, M. (Eds.): Countryside Stewardship: Farmers, Policies and Markets. Elsevier Science, Oxford, UK
13. Suter II, G. W. (1999): Developing Conceptual Models for Complex Ecological Risk Assessments. Human and Ecological Risk Assessments 5, 375–396
14. U.S. Environmental Protection Agency (1998): Guidelines for Ecological Risk Assessment. Washington, D.C.

Improve the Shipping Performance of Build-to-Order (BTO) Product in Semiconductor Wafer Manufacturing

Shao-Chung Hsu, Chen-Yuan Peng and Chia-Hung Wu

Macronix International Co., Ltd., No. 16, Li-Hsin Road, Science-Based Industrial Park, Hsinchu, Taiwan, R.O.C

Abstract

The aim of this paper is to establish an application to review and control the shipping performance of build-to-order (BTO) product in semiconductor wafer manufacturing. The final goals are to avoid shortage of each order released from wafer bank to fulfill the finished goods requirement and minimize the overage quantity that will induce the high inventory. We apply the statistical method in wafer sort yield prediction via history data. Base on this, we can model
The overage and shortage quantity with the influential variables, namely order size, yield variability, uncertain demand and customer acceptance criterion. The model also allows estimating the expected overage as the review criterion between manufacturing unit and sales unit.
Keywords: BTO, Shortage, Overage, Yield Distribution

1. Introduction

The core competency of semiconductor manufacturing is composed by the high yield, the short manufacture leading time and the low unit cost. There are two planning strategies for product production. One is the build-to-stock (BTS) or make-to-stock (MTS). The other is the build-to-order (BTO) or make-to-order (MTO) [1]. The BTS strategy is to produce the finish product according to the production plan. The BTO strategy is to produce the finish good until receiving the order from customers.

The unit during the process in semiconductor wafer manufactory is the wafer. Several wafers will be grouped into a lot during the process. But the order taken from customer is counted via dies for BTO product. So the order quantity in die must be changed into equivalent wafers before manufacturing. The equivalent wafers can be transferred by gross die, yield and order quantity in formula (1).

$$\text{Released Wafers} = \frac{\text{Order Quantity by Die}}{\text{Gross Die} * \text{Yield}} \quad (1)$$

The yield in the formula (1) includes manufacturing line yield, circuit probe (CP) yield, assembly yield and final test yield. The CP yield has the largest variation than the others. So we will only focus on the impact from CP yield in this paper.

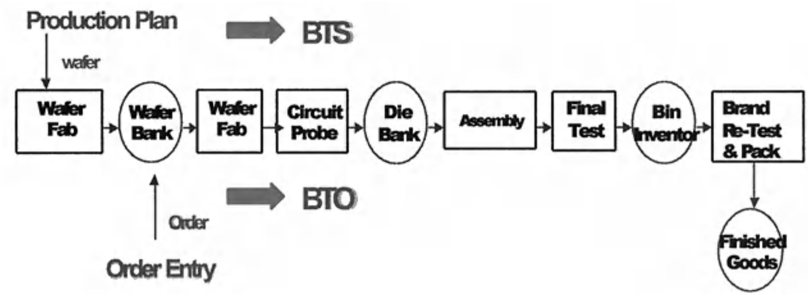


Fig. 1. Semiconductor Manufacturing Process with BTO and BTS strategy

The result of each delivered order is either shortage or overage. It depends on the difference of forecasted yield (Y_p) and the actual yield (Y_A). If Y_A is less than Y_p, then the order will be shortage. Otherwise, it will be overage. For guarantee the order on time and on quantity to deliver to customer, the wafer manufactory would rather produce the overage quantity to avoid shortage. But the problem of the overage quantity is usually ignored since there is no direct pressure from the customer. The high overage quantity will induce the high inventory. This kind of inventory usually is the dead inventory and induces the lost in finance if no repeated order entry in the future.

The BTO product in other industry, such as computer or the main board assembly, has defined the fixed discount rate between the wafer manufactory and customer. The major inventory of these industries is come from the uncertainty demand quantity which is caused by the customer committed quantity and the order entry time [2-7]. But for the semiconductor manufacturing, the characteristic of BTO product is that the wafer, in terminate part, will be divided into hundreds of final product (die). The change rate is depended on the actual yield, which is a random variable with statistical distribution. So, the inventory of this kind of product is caused by the combination of uncertain supply and the uncertain demand.

The aim of this paper is to develop a yield forecast model. With this model, we can develop the application to avoid shortage of each order released from wafer bank to fulfill the order and minimize the overage quantity.

2. The Yield Forecast Model

There are two methods to forecast the CP yield in semiconductor manufacturing. One is calculated from the defect data, which is collected during manufacturing [8]. The other is to apply statistical method to the yield distribution. In this paper, we will focus on the distribution model. There are two assumptions to apply in this model. The first is the yield of lot is a random variable from normal distribution with mean (μ) and standard deviation (s). The second assumption is each lot assigned to manufacture is of full size with 25 wafers. The number of lot for each order can be derived from the requested wafer via the second assumption.

According to the central limit theorem [9], we can derive the yield distribution of order from the yield distribution by lot and the released lot quantity. The mean of yield distribution by order will not change from the different order size by lot, but the variance will change with the number of lot.

To guarantee the on time delivery rate and reduce the shortage by order, the yield must be forecasted lower than μ -k*s , where k can be adjusted by the customer grade level. The customer grade level is depended on the priority of order and the importance of customer. For example, the manufactory can divide the customers into 3 grades: N, A and B. The on time delivery rate can be defined as 95%, 85% and 66%, respectively. The corresponding k is 1.645, 1.035 and 0.385, which can be got from normal distribution table. The overage of order can be defined as formula (2):

$$Overage\left(Y_A\right) = W * GSD * \left(Y_A - Y_p\right) \tag{2}$$

where

W : the number of wafers released to manufactur e

GSD : the gross die of product

We can assume the Y_A comes from the normal distribution then derive the expected overage from formula (3). The result shows that the factors impacted the overage is the released lot quantity of order, the standard deviation of product yield, the tolerance of on time delivery rate of the customer grade and the gross die of product. We can also define the performance metrics of overage, OO ratio, by formula (4). This index, OO ratio, can indicate the performance of each order or by product.

$$E(Overage) = \int_{Yp}^{\infty} Overage(Y_A) f(Y_A) dY \tag{3}$$

$$= W * GSD * \left\{ s * \left(\frac{1}{\sqrt{2p}} * e^{-(\frac{k^2}{2})} + k * f \ (k) \right) \right\}$$

$$= W * GSD * s * f(k)$$

where

$$Y_P = \mu - k * s$$

$$k = \frac{(\mu - Y_P)}{s}$$

φ : the standard normal probability desnity funtion

f(k) :the function related to order size

$$\text{OO ratio} = \frac{\text{Overage Quantityby Die}}{\text{Order Qantityby Die}} * 100\% \tag{4}$$

3. Computational Simulation

The overage model and performance metric can be verified by computational simulation. The conditions of the simulation are: the customer grade is N grade with acceptance delivery level 95% and the order size is one lot with 25 wafers. We change the yield distributions with different mean and standard deviation. The result is as Table (1). From Table (1), we can find that the forecasted yield (Yp) must be forecasted lower to meet the acceptance delivery level if the standard deviation (std) is large. But it will be suffered the increasing of overage quantity and higher the OO ratio.

Table 1. The simulation Result (1)

Mean	std	Y_p	OO ratio
0.8	0.02	0.7671	4.3%
0.8	0.05	0.7178	11.6%
0.8	0.1	0.6355	26.2%
0.7	0.05	0.6178	13.5%
0.8	0.05	0.7178	11.6%
0.9	0.05	0.8178	10.2%

We can apply another simulation to check the impact of order size to overage. We use the same conditions of simulation of last simulation and fix the yield mean to 0.8. We change the order size from 1 lot to 10 lots and change the standard deviation. The final result is as Fig. 2. We can find that the OO ratio will be small if the order size is large. . For example, if we want to control OO ratio under 5%, then the required order size for Product B needs release 6 lots and for Product C needs 10 lots. Since the standard deviation of Product D is large, it is hard to control the OO ratio under 5%. This means the standard deviation of yield is the key factor for overage control.

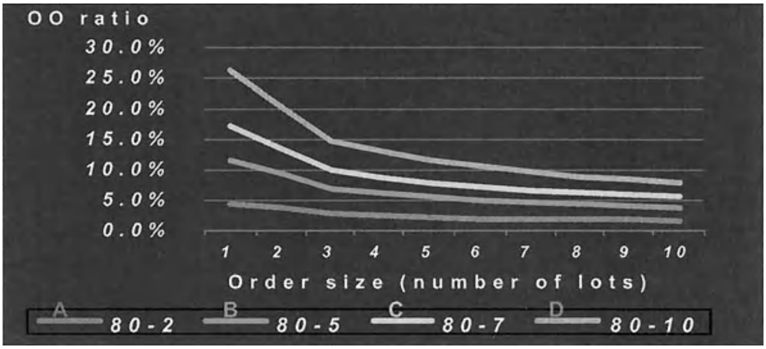


Fig. 2. The simulation result (2)

4. An Empirical Case and the Application

The data of this empirical case is come from a semiconductor company located in Taiwan. There are several products in this company. We can examine the actual OO ratio with the order size and standard deviation in Table 2. The result is consistent with the result of simulation. There are several applications in the system to assist the production control engineers to control the inventory and cost (Fig. 3).

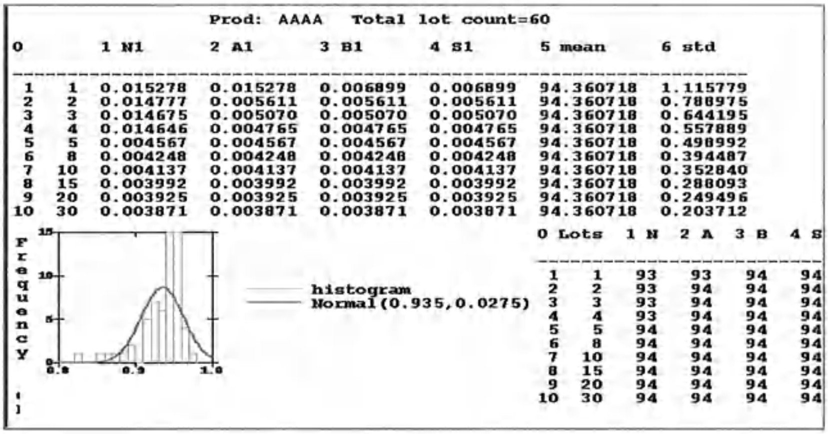
Table 2. : An Empirical Case of the Overage Model

Tech	Prod	Mean	Std	Actual OO ratio		
				N	A	B
M25	A	94.36	1.12	1.5%	1.5%	0.7%
	B	91.09	2.35	4.7%	2.6%	1.8%

C	88.37	3.46	6.6%	4.3%	2.5%
D	86.11	3.44	7.7%	4.2%	2.3%
E	85.29	3.21	6.7%	4.3%	2.4%
F	83.21	5.97	14.1%	8.7%	4.5%
G	81.07	3.03	6.7%	4.3%	2.3%
H	79.80	5.25	12.0%	7.8%	4.5%
I	72.57	7.30	19.2%	12.5%	6.2%

5. Conclusion and Future Work

In this paper, we define the expected overage is function of the order size and customer grade. We also distinguish the overage amount by the manufacturing and non-manufacturing factors.From the model, we can conclude the standard deviation is the key factor for the overage of BTO product. The product with unstable yield (large standard deviation) will increase the overage quantity and OO ratio when the order size is small. At last ,The empirical yield distribution is always skew with lorng tail in the real environment..This violates the assumption of normal distribution We can enhance the methodology by changing from normal to skew distribution in the future.



Prod: AAAA Total lot count=60

0		1 N1	2 A1	3 B1	4 S1	5 mean	6 std
1	1	0.015278	0.015278	0.006899	0.006899	94.360718	1.115779
2	2	0.014777	0.005611	0.005611	0.005611	94.360718	0.788975
3	3	0.014675	0.005070	0.005070	0.005070	94.360718	0.644195
4	4	0.014646	0.004765	0.004765	0.004765	94.360718	0.557889
5	5	0.004567	0.004567	0.004567	0.004567	94.360718	0.498992
6	8	0.004248	0.004248	0.004248	0.004248	94.360718	0.394487
7	10	0.004137	0.004137	0.004137	0.004137	94.360718	0.352840
8	15	0.003992	0.003992	0.003992	0.003992	94.360718	0.288093
9	20	0.003925	0.003925	0.003925	0.003925	94.360718	0.249496
10	30	0.003871	0.003871	0.003871	0.003871	94.360718	0.203712

0	Lots	1 N	2 A	3 B	4 S
1	1	93	93	94	94
2	2	93	94	94	94
3	3	93	94	94	94
4	4	93	94	94	94
5	5	94	94	94	94
6	8	94	94	94	94
7	10	94	94	94	94
8	15	94	94	94	94
9	20	94	94	94	94
10	30	94	94	94	94

Fig. 3. The output of the order release system to estimate the expected overage

References

[1] Krajewski, L.J. and Ritzman, L.P., Operation Management: Strategy and Analysis, Addison-Wesley Longman Inc. (2000)
[2] Erdema, A.S. and Özekici, S., Models with Random Yield in a Random Environment, Int. J. Production Economics 78 (2002), 239–253
[3] Gerchak, Y., Vickson, R.G. and Parlar, M., Periodic Review Production Models with Variable Yield and Uncertain Demand, IIE Transactions 20 (1988), 144-150
[4] Parlar, M. and Perry, D., Analysis of a (Q,r,T) Inventory Policy with Deterministic and Random Yields When Further Supply is Uncertain, European Journal of Operation Research (1995) 84, 431~443.
[5] Parlar, M and Wang, D., Diversi?cation under yield randomness in inventory models, European Journal of Operational Research 66 (1993), 52–64
[6] Sepheri, M., Silver, E.A. and New, C., A Heuristic for Multiple Lot Sizing for an Order Under Variable Yield, IIE Transactions 18 (1986), 63-69
[7] Yano, C.A. and Lee, H.L., Lot sizing with random yields: A review, Operations Research 43 (2) (1995), 311–334
[8] Ferris-Prabhu, A.V., VLSI Yield Management and Prediction, Quality and Reliability Engineering International Vol. 1 (1985), 219-225
[9] Bhattacharyya, G.K. and Johnson, R.A., Statistical Concepts and Methods, John Willey & Sons (1997)

Competence Set Expansion for Obtaining Scheduling Plans in Intelligent Transportation Security Systems

Yi-Chung Hu[1], Yu-Jing Chiu[2], Chin-Mi Chen[3], Gwo-Hshiung Tzeng[2]
[1]*Institute of Information Management, National Chiao Tung University, Hsinchu, Taiwan, ROC*
[2]*Institute of Management of Technology, National Chiao Tung University, Hsinchu, Taiwan, ROC*
[3]*School of Nursing, National Defense Medical Center, Taipei, Taiwan, ROC*

Abstract

In this paper, to reach safe driving, a new method is proposed to acquire project schedules by competence set expansion for developing intelligent transportation systems (ITSS), that can promote security, efficiency and comfort for drivers. Since for each decision problem, there is a competence set consisting of ideas, knowledge, information, and skills needed to solve that problem, this paper treats ITSS as the needed competence set to attain the goal of safe driving. Schedules can be further obtained by using a known method proposed by Li (1999) for expanding competence sets. An empirical study is utilized to demonstrate the usefulness of the proposed method.

Keywords: Competence set; Scheduling; Intelligent transportation systems; Integer programming.

1. Introduction

Intelligent transportation systems (ITS) aim to provide improvements to the efficiency and safety of transportation systems by applying new information and communication technologies (Berbineau, 1999). ITS can also provide real-time information to help drivers or travelers avoid possible traffic problems, such as traffic congestion. With the growing awareness of the potential importance of ITS, the governments of many country have paid more attention on the development and the academic studies on ITS.

Several kinds of necessary information can be provided to drivers while they are driving, such as rescue services or emergency services (Iguchi, 2002). However, the traffic information services should be primarily aimed for safe driving (Iguchi, 2002). In the statistical data reported by the US Department of Transportation Federal Highway Administration (1998, 1999), it can be found that the effects of intelligent transportation security systems (ITSS) on the traffic transportation were significant. For reaching the goal of safe driving, the acquisition of schedules in the project with respect to the development of ITSS is significant. ITSS project can be divided into a number of activities for developing security information subsystems. However, it is possible that the exact order among the activities is not easily determined.

Competence sets were initiated by Yu (1990). Its mathematical foundations were further established by Yu and Zhang (1990). It is considered that for each decision problem there exists a competence set consisting of ideas, knowledge, information, and skills for solving that problem. When decision makers have acquired the needed competence set and are proficient in it, they will be comfortable and confident in making decisions (Li, Chiang, & Yu, 2000). In order to acquire a needed competence set to make a decision, finding appropriate learning sequences to acquire single skills needed for decision makers, the so-called competence set expansion, is necessary. From these viewpoints, it is reasonable that we treat a set of undeveloped subsystems in ITSS as a needed competence set for attaining the goal of safe driving. If each subsystem is further treated as a needed skill, then the competence set expansion can serve as a scheduling plan for undeveloped subsystems.

On the other hand, it is also known that learning directly from one skill to another skill requires learning cost. Generally speaking, the stronger the relationship that exists between two skills, the smaller is the learning cost between these two skills (Hu, Chen, Tzeng, & Chiu, 2002). We thus propose a relationship-based method based on the grey relational analysis (Deng, 1982) to determine learning costs. To effectively facilitate the promotion of ITSS, the main aim of this paper is to provide a new method that can acquire the appropriate schedules of undeveloped subsystems through the proposed relationship-based method and a known integer programming method proposed by Li (1999) for effectively expanding the competence set.

The rest of this paper is organized as follows. At first, we introduce the concepts of the competence set expansion in Section 2. A relationship-based method used for determining the learning costs is proposed in Section 3. Section 4 briefly describes the Li's integer programming method. A real case for demonstrating the feasibility of facilitating the development of several ITSS subsystems is demonstrated in Section 5. We end this paper with conclusions in Section 6.

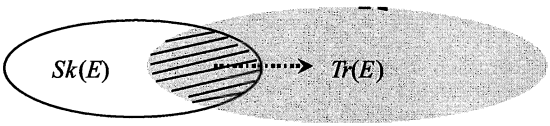


Fig 1. Competence set expansion.

2. Competence Set Expansion

For each decision problem E, there is a competence set, denoted by $CS(E)$, consisting of ideas, knowledge, information, and skills for successfully solving that problem. In addition, there exists a skill set denoted by $Sk(E)$ that has been acquired by decision makers, and a truly needed competence set denoted by $Tr(E)$. Decision makers must acquire $Tr(E)\backslash Sk(E)$ from the existing competence set (i.e., $Sk(E)$) through the competence set expansion to resolve E.

A competence set expansion represents the way to find an effective way to generate learning sequences by acquiring the skills which are truly needed so that $Tr(E)\backslash Sk(E)$ can be obtained (Feng & Yu, 1998). We depict the concept of the competence set expansion in Fig. 1, where the shaded area containing no any lines is $Tr(E)\backslash Sk(E)$. When decision makers have not acquired $Tr(E)\backslash Sk(E)$, it is more difficult for them to make decisions. In this paper, for simplicity we assume that $Sk(E)$ is an empty set.

We consider that if the relationship that exists between two single skills, say f_1 and f_3, is much stronger than that between another two skills, say f_2 and f_3, then it is more practical to acquire f_3 from f_1 instead of from f_2. We further interpret the learning cost, say $c(f_i,f_j)$, to be promotion cost for acquiring f_j from f_i. In fact, it seems to be impossible to exactly measure learning costs in advance using either money or time.

3. A Relationship-Based Method

A relationship-based method based on the grey relational analysis is proposed for finding learning costs. Actually, there exist distinct relationships between any two subsystems in the real world (Deng, 1982). Grey theory, as proposed by Deng (1982), can perform grey relational analysis for these subsystems by dealing with finite and incomplete output data series obtained from these subsystems (Huang & Huang, 1997).

We treat each ITSS subsystem as a needed skill f_p, and its finite output data series is $(f_{p_1}, f_{p_2}, \ldots, f_{p_n})$ where f_{p_i} $(0 \leq f_{p_i} \leq 1)$ is the part-worth of the i-th criterion with respect to f_p. That is, we assume that each subsystem can be evaluated by n various criteria. Then, the grey relations can be employed to find the learning cost between any two skills. Generally speaking, the larger the grey relation that exists between two skills, say f_p and f_i, the smaller is learning cost between f_p and f_i.

Given one reference sequence, say the p-th single skill f_p ($1 \leq p \leq K$, where K is the number of single skills), and some comparative sequences, say the i-th single skill f_i ($1 \leq i \leq K$), we can easily obtain the grey relation between f_p and f_i by viewing f_p as a desired goal. Formally, given the reference sequence f_p and the comparative sequences f_i with the normalized form, the grey relational coefficient (GRC) $\xi(f_{p_j}, f_{i_j})$ between f_{p_j} and f_{i_j} $(1 \leq j \leq n)$ can be computed as Eq. (1) (Huang & Huang, 1997; Hsu & Chen, 2000).

$$\xi(f_{i_j}, f_{p_j}) = \frac{\Delta_{\min} + \rho \Delta_{\max}}{\Delta_{ij} + \rho \Delta_{\max}} \tag{1}$$

where ρ is the discriminative coefficient $(0 \leq \rho \leq 1)$, and usually $\rho = 0.5$ (Hsu & Chen, 2000). Moreover,

$$\Delta_{\min} = \min_i \min_j |f_{p_j} - f_{i_j}|, 1 \leq i \leq K, 1 \leq j \leq n \tag{2}$$

$$\Delta_{\max} = \max_i \max_j |f_{p_j} - f_{i_j}|, 1 \leq i \leq K, 1 \leq j \leq n \tag{3}$$

$$\Delta_{ij} = |f_{p_j} - f_{i_j}| \tag{4}$$

where $|\,|$ denotes the absolute value. Clearly, $\xi(f_{i_j}, f_{p_j})$ is between zero and one. Then, the grey relational grade (GRG) denoted by $\Upsilon(f_i, f_p)$ can be computed as Eq. (5).

$$\Upsilon(f_i, f_p) = \frac{1}{n}\sum_{j=1}^{n} \xi(f_{i_j}, f_{p_j}) \tag{5}$$

Thus holds $0 \leq \Upsilon(f_i, f_p) \leq 1$, and the larger the value of $\Upsilon(f_i, f_p)$, the closer the relationship is between f_p and f_i. The learning cost for directly learning f_p from f_i, denoted by $c(f_i, f_p)$, is heuristically computed as Eq. (6).

$$c(f_i, f_p) = 1 - \Upsilon(f_i, f_p),\ 1 \leq i, p \leq K \tag{6}$$

It is clear that $0 \leq c(f_i, f_p) \leq 1$ also holds here. Eq. (6) indicates the relationship between the learning cost (i.e., $c(f_i, f_p)$) and the grade of relationship (i.e., $\Upsilon(f_i, f_p)$). A learning cost table can be thus built. As we have mentioned above, it is also reasonable that $c(f_i, f_p)$ is interpreted as the cost for directly promoting f_p from f_i. It is noted that $c(f_p, f_p)$ $(1 \leq p \leq K)$ does not exist.

4. Generate Learning Sequences

Below, we briefly describe a useful integer programming method proposed by Li (1999), which is used to effectively generate learning sequences with minimum learning cost. In fact, a competence set expansion can also be roughly regarded as a spanning tree construction process (Feng & Yu, 1998). Moreover, the arc in the tree is directed. Additionally, all nodes representing single skills must take part in the generation of the final learning sequence.

Actually, all skills construct a digraph G consisting of subgraphs S and T, and each skill corresponds to a node. S consists of skills in $Sk(E)$, and T consists of skills in $Tr(E)\backslash Sk(E)$. The learning sequences are generated for $Tr(E)\backslash Sk(E)$ starting from $Sk(E)$. For G, several definitions must be given. First, define $|A(i)|$ and $|B(i)|$ as the numbers of nodes immediately before node i and immediately after node i, respectively. In addition, define u_i for node i as $u_i = 1$ if node i takes part in the generation of the learning sequence; otherwise $u_i = 0$. Also, define $v(i, j)$ for the arc connecting node i to node j as $v(i, j) = 1$ if $v(i,j)$ is one path of the learning sequence; otherwise $v(i, j) = 0$. Let $V(S)$ and $V(T)$ be sets of nodes of S and T, respectively, then both u_i and $v(i, j)$ are 0-1 variables that satisfy the following properties:

a. $u_i = 1$, for each $i \in V(S)$ or i corresponding to a single skill in T (7)

b. $|A(i)|\, u_i \geq \sum_{j\in A(i)} v(i,j)$, if $i \in V(S)$ (8)

c. $u_i \leq \sum_{j\in B(i)} v(j,i)$, if $i \in V(T)$ and i is not a compound node (9)

d. $(|A(i)| + |B(i)|)\, u_i \geq \sum_{j\in B(i)} v(j,i) + \sum_{j\in A(i)} v(i,j)$, if $i \in V(T)$ (10)

In addition, let λ_i be the sequence number of node i in the learning sequence; then the following relations hold:

a. $\lambda_i = 0$, if $i \in V(S)$ (11)

b. $\lambda_i - \lambda_j + K\, v(i,j) \leq K - 1$ (12)

c. $u_i \leq \lambda_i \leq K\, u_i$ (13)

where λ_i is an integer variable. Eqs. (7)-(10) find those arcs which can be contained in the learning sequence, and Eqs. (11)-(13) find the sequence number of each node in the learning sequence. The more details on the above-mentioned properties can be found in (Li, 1999).

Since $Sk(E)$ is assumed to be an empty set, several revisions of the Li's method must be made. Let S consist of a virtual node labeled by 0 (i.e., $V(S) = \{0\}$). Furthermore, the virtual node (i.e., node 0) is directly linked to each node i by a directed arc with learning cost being equal to zero, where node $i \in T$. Our purpose is to find the starting node labeled by $n_0^{(T)}$ (i.e., $n_0^{(T)} \in V(T)$) immediately after node 0 in the learning sequence. If nodes $i_1, i_2,\ldots, i_K$ represent single skills $f_1, f_2,\ldots, f_K$, respectively, then the following relation holds:

$$\sum_{j=1}^{K} v(0,i_j) = 1 \tag{14}$$

An integer program can be further formulated by combining Eqs. (7)-(14) and giving an object function as follows:

$$\text{Minimize cost(G)} = \sum_{r(i,j)\in E} c(i,j)\cdot v(i,j) \tag{15}$$

where E is the set of arcs of G. That is, the objective is to find a spanning tree with minimum learning cost in G (i.e., minimum spanning tree).

5. Empirical Results

The development of fourteen undeveloped subsystems in ITSS, as suggested in the technical reports of the Taiwan Ministry of Transportation and Communications (1998, 1999), and from opinions of ITS project managers in the official or academic institutions, such as the Institute of Transportation of the Taiwan Ministry of Transportation and Communications, are considered in this section. These subsystems should be completed in the near future.

The basic functions of the undeveloped subsystems include information on weather and on road status, intelligent navigation, status examination of the vehicle, with warning for hazards (f_4), warning for the dangerous status of the vehicle, warning for the driver's physical conditions and abnormal operations (f_6), assistance in driving (f_7), automatic extinguishing fire (f_8), air bags (f_9), the notification of an accident (f_{10}), absorption of collision and deceleration speed of the vehicle (f_{11}), recording of driving status (f_{12}), automatic relief of locked status (f_{13}), and intelligent tires (f_{14}).

Additionally, seventeen criteria with equal weight (i.e., 1/17) are used to evaluate each subsystem, including easy use or control, easy maintenance of equipments (c_2), and high protection (c_3), and so on. The part-worth of each criterion's level is obtained using the questionnaire asking decision makers to pick a statement that best describes the given criterion's level (Yoon & Hwang, 1995). Actually, each criterion is evaluated on a five-point scale. That is, 'very insignificant' is evaluated with part-worth 1 and 'insignificant' with part-worth 2, and so on. Considering response quality, the 10 professional respondents were chosen from ITS project managers in the government or academic institutions. The learning cost table can be thus obtained by Eqs. (5) and (6), and is omitted for simplicity.

Let f_i be the i-th node ($1 \leq i \leq 14$) in G. Then, a mathematical integer program to find a learning sequence with minimum learning cost can be formulated by Li's method. The detailed program is omitted for simplicity. Then solve this mathematical program by the LINGO package to obtain the optimal solution. The sequence depicted in Fig. 2 serves as a scheduling plan to develop the subsystems, providing relative orders of construction or promotion of these subsystems are obtained. In addition, the scheduling of the fourteen subsystems of ITSS is also easily described. For example, we can see that f_2 is suggested for decision makers to construct first among the fourteen subsystems. Subsequently, f_4 and f_{11} can be developed simultaneously when f_8 has been constructed.

6. Conclusions

A new method is proposed to acquire appropriate schedules for undeveloped subsystems through two primary steps: one, to find learning costs between any two subsystems in ITSS by using the proposed relationship-based method derived from grey relational analysis; and the other, to generate a schedule with the minimum learning or promotion cost. The main characteristic of the proposed method is to pertinently treat a set of undeveloped subsystems in ITSS as a needed competence set for attaining the goal of safe driving. Additionally, it seems that the proposed method can be utilized to evaluate the development order of other types of information systems.

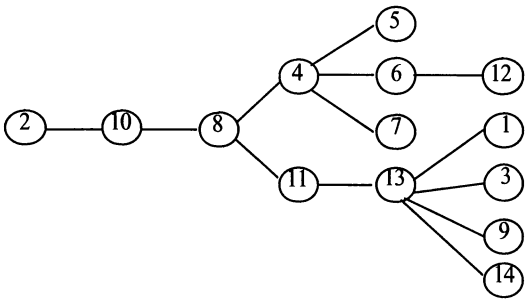


Fig. 2. Scheduling in ITSS.

References

Berbineau, M. (1999). Track-to-train transmission systems: State of the art and outlook. *Recherche Transports Securite*, 62, 35-48.

Deng, J. L. (1982). Control problems of grey systems. *Systems and Control Letters*, 1(5), 288-294.

Feng, J. W., & Yu, P. L. (1998) Minimum spanning table and optimal expansion of competence set. *Journal of Optimization Theory and Application*, 99(3), 655-679.

Hsu, Y. T., & Chen, C. M. (2000). A novel fuzzy logic system based on N-version programming. *IEEE Transactions on Fuzzy Systems*, 8(2), 155-170.

Huang, Y. P., & Huang, C. H. (1997). Real-valued genetic algorithms for fuzzy grey prediction system. *Fuzzy Sets and Systems*, 87(3), 265-276.

Institute of Transportation, Taiwan Ministry of Transportation and Communications (1998). The Progress and Relative Techniques in Intelligent Transportation Systems. *Technical report*, Taiwan: Ministry of Transportation and Communications.

Institute of Transportation, Taiwan Ministry of Transportation and Communications (1999). The Development Outline of Intelligent Transportation Systems in Taiwan-The Investigation of The Application Domains and The Supply and Demand of User Services. *Technical report*, Taiwan: Ministry of Transportation and Communications.

Li, J. M., Chiang, C. I., & Yu, P. L. (2000). Optimal multiple stage expansion of competence set. *European Journal of Operational Research*, 120(3), 511-524.

Li, H. L. (1999). Incorporating competence sets of decision makers by deduction graphs. *Operations Research*, 47(2), 209-220.

US Department of Transportation Federal Highway Administration (1998). Transportation Planning and Intelligent Transportation Systems: Putting the Pieces Together. *Technical report*, USA Department of Transportation Federal Highway Administration.

US Department of Transportation Federal Highway Administration (1999). Intelligent Transportation Systems Benefit. *Technical report*, USA Department of Transportation Federal Highway Administration.

Yoon, K. P., & Hwang, C. L. (1995). *Multiple Attribute Decision Making: An Introduction*, London: Sage Publications.

Yu, P. L. (1990). *Forming Winning Strategies: An Integrated Theory of Habitual Domain*, New York: Springer-Verlag.

Yu, P. L., & Zhang, D. (1990). A foundation for competence set analysis. *Mathematical Social Sciences*, 20, 251-299.

DEA for Evaluating the Current-period and Cross-period Efficiency of Taiwan's Upgraded Technical Institutes

Li-Chen Liu, Chuan Lee, and Gwo-Hshiung Tzeng*

Graduate Institute of Management Science, Ming Chuan University, Taipei, Taiwan; *Institute of Management of Technology, National Chiao Tung University, Hsinchu, Taiwan; e-mail: ghtzeng@cc.nctu.edu.tw

Abstract. *This study used Data Envelopment Analysis (DEA) to examine the relative managerial efficiency for evaluating current-period and cross-period efficiency of 38 technological institutes upgraded from junior colleges in Taiwan in 1998. In addition, the managerial efficiency variations of each individual institute in between 1995 and 1998 were also determined. The study results show that the operational category is significant among primary analysis variants, in other words, private schools perform significantly better than public schools in terms of managerial efficiency. Furthermore, school size is significant, with schools having more than 201 classes achieving higher managerial efficiency. However, geographical location is not significant. This study also verified that integration of the results of both relative managerial efficiency analysis and managerial efficiency variation analysis could be a powerful approach to help design managerial strategies that are both appropriate and feasible.*

Keywords: data envelopment analysis (DEA), managerial efficiency, performance evaluation, technological institute

1. Introduction

Most of those previous studies evaluating higher education quality employed holistic performance indices, which had a significant weakness in that very few of them considered the weighting distribution for each criteria performance (Higgins, 1989). Bates (1997) employed Data Envelopment Analysis (DEA) technique to study the relative efficiency of the Great Britain education enterprise from a socio-economic aspect. In their study on the evaluation of school managerial efficiency and curriculum planning, Bessent et al. (1983) employed the DEA technique and suggested some input and output weightings. Kao (1994) used the Pareto Optimality theory to survey the performance of Taiwan's junior colleges and to sort their grade rankings, yielding outcomes that met those of the official study by the Education Ministry of Taiwan. As compared to conventional evaluation approaches, it is possible for the DEA technique to produce more reasonable and accurate outcomes, and some studies have supported this conclusion, such as Charnes et al. (1981) and Ahn et al. (1989). In the published studies, to administer a school performance evaluation, researchers commonly employed field visits, although they did not consider the individual differences of those schools struggling to be upgraded from junior colleges. This researcher proposes that a critical study should examine the differentiations among governing variables, such as geographic environment, educational resources, developmental potentials, and school size. This specific concern is the goal of this paper. In an attempt to specify strategies that are more reasonable and feasible for increasing school efficiencies, there is a need for a new evaluation model that is more careful and objective as well as designing related administration process. This model should assess the reasonability of both resource inputs and outputs, and provide decision makers sufficient information to formulate more appropriate educational plans. Therefore, this study measured the managerial efficiencies of 38 of Taiwan's technological institutes, examing the efficiency of inputs and outputs. First evaluation indices are constructed by means of group brainstorming, literature investigation, and interviews with experts. Then, the relative managerial efficiencies of each individual target institute for the year 1998 were evaluated, as were the managerial efficiency variations between 1995 and 1998, by employing the cross-period DEA approach. Study results are listed as the following: (1) outcomes of relative efficiency show significant differences in which school managerial categories and school size; (2) there are not significant differences due to geographic location of schools. Analysis results of cross-period efficiency variations show that both Malmquist productivity indices and technical change (*TC*) are declining, and that technical efficiency change (*EC*) is increasing. This study should provide site management with some improvement directions and governing indices for both operational management and resource application, and also provide educational officials and experts with a appropriate evaluation system.

The remainder of this paper is organized as follows. Section 2 explains the selection of input and output variables; the section 3 states the static and dynamic evaluations of the DEA approach; section 4 includes the empirical analysis and related discussions; and the last section provides discussion and recommendations.

2. The Selection of School Objects and Variables for Performance Evaluation

This study use the DEA technique to measure school efficiency. Original data come from both input and output data of each decision making unit (DMU) of 38 technological institutes. Determining the relationship between organizational goals as well as both input and output is the goal of this study.

2.1 Research samples

This study procedure was two-stepwise: firstly, the researchers collected reachable data of these 38 technological institutes for empirical study, then administering a series of efficiency evaluations; secondly, the researchers analyzed the cross-period efficiency variation of these technological institutes between 1995 and 1998.

2.2 Selection of both Input and Output Indices

Those suggested indices of both input and output, collected from related studies, are listed as the following: educational resources as well as the quality of faculty and students (Wang, 1993 & 1995); the unit cost per student (Power, 1989; Chang, 1995); the number of full-time faculty (Hufner, 1987; Chang, 1995); faculty-to-student ratio (Hufner, 1987; Power, 1989); general affairs and managerial expenditures (Ahn et al., 1989); and the expenditures of constant properties (Ahn et al., 1989). The suggested output indices include: the achievement of instruction, research, and service (Chang, 1995); the case numbers and budgets of contracted projects (Hufner, 1987; Ahn et al, 1989); as well as the number of graduates (Inbar, 1988). The input and output indices of this study came from the synthesis of those suggested in the related studies above, and from the concerns of the feasibility in practical administration and data collection. These input variables have four dimensions: physical resources, human resources, hardware resources, and information resources. The five input indices include: building area, numbers of assistant professors and higher, annual expenditures, size of library collection, and numbers of periodicals. Output variables are three dimensional, teaching efficiency, practical research efficiency, and service efficiency. The output indices are: the number of graduates, research expenditures, and incomes from both school-industry-collaboration and continuing education.

3. Building the Performance Model

This study employs the DEA technique, presented by Charnes et al. (1978), to do an analysis of the same kind.

3.1 Classical Radial Efficiency Measure

According to the CCR model, presented by Charnes, et al. (1978):

$$\text{Max } h_a = \frac{\sum_{r=1}^{S} u_r y_{ra}}{\sum_{i=1}^{m} v_i x_{ia}} \tag{1}$$

subject to:

$$\frac{\sum_{r=1}^{S} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} \le 1, \quad k = 1,2,\cdots,n$$

$$0 < e \le u_r, 0 < e \le v_i, \quad i = 1,\cdots,m, \quad r = 1,\cdots,s$$

where x_{ik} stands for the i^{th} input value of the k^{th} DMU; y_{rk} stands for the r^{th} output value of the k^{th} DMU; u_r and v_i stand for the weight of the r^{th} output and i^{th} input respectively, e is a non-archimedean quantity which is usually given a small possitive value, for instance 10^{-5} ; h_a is relative efficiency value.

Charnes et al. (1978) transformed the CCR model from a fractional programming model into a linear programming model in a major attempt to improve the operational convenience, necessary because the conventional CCR model is subjected to numerous constraints amid the searching of solutions.

$$\text{Max } h_a = \sum_{i=1}^{s} u_r \, y_{ra} \tag{2}$$

subject to:

$$\sum_{i=1}^{m} v_i x_{ik} - \sum_{r=1}^{s} u_r y_{rk} \geq 0; \; k = 1,\cdots,n$$

$$\sum_{i=1}^{m} v_i x_{ia} = 1$$

$$u_r \geq \varepsilon > 0, \;\; r = 1,\cdots,s; \qquad v_i \geq \varepsilon > 0, \;\; i = 1,\cdots,m$$

The number of elements in the dual problem in Eq. (2) can be reduced for the sake of simplicity and the dual problem after conversion becomes:

$$\text{Min } \{ \; ?_a - e \; [\; \sum_{i=1}^{m} S_{ia}^- + \sum_{r=1}^{s} S_{ra}^+] \; \} \tag{3}$$

subject to:

$$?_a x_{ia} - \sum_{k=1}^{n} ?_k x_{ik} - S_{ia}^- = 0 \quad i = 1,\cdots,m$$

$$y_{ra} - \sum_{k=1}^{n} ?_k y_{rk} + S_{ra}^+ = 0; \;\; r = 1,\cdots,s$$

$$S_{ia}^-, S_{ra}^+, \lambda_k \geq 0$$

where s_{ia}^- and s_{ra}^+ are slack variables. The dual problem, presented by Banker et al. (1986), has two primary strengths, the reduction of calculation barriers and the provision of more helpful information. When an individual DUM achieve Pareto's optimality situation $\theta_a^* = 1$ ("*" stands for the optimal solution) i. e.,

$\{(x_{ia}^*, y_{ra}^*) | S_{ia}^- = S_{ra}^+ = 0, \;\; i = 1,\cdots, m; \;\; r = 1,\cdots,s\}$, and $S_{ia}^- = S_{ra}^+ = 0$ When a DMU has not achieve Pareto's optimality situation, we can make some adjustments by using Eq.(4), and then helping individual DMU to achieve Pareto's optimality situation (relative efficientness).

$$\begin{aligned} x_{ia}^* &= \theta^* x_{ia} - S_{ia}^{-*}, i = 1,\cdots,m \\ y_{ra}^* &= y_{ra} + S_{ra}^{+*}, r = 1,\cdots,s \end{aligned} \tag{4}$$

3.2 Cross-period Efficiency Analysis – Testing Malmquist Index

According to Fare et al. (1992), the Malmquist productivity index is a product of technical change (*TC*) and efficiency change (*EC*). Eqs. (5), (6), and (7) denote *TC*, *EC*, and the Malmquist productivity index respectively.

(a) Technical Change (TC) is denoted by Eq. (5). When TC > 1, technical gress is indicated, on the contrary, TC < 1 indicates technical regress.

$$TC=\left[\frac{D^{t+1}(X^{t+1},Y^{t+1})}{D^{t}(X^{t+1},Y^{t+1})}\;\frac{D^{t+1}(X^{t},Y^{t})}{D^{t}(X^{t},Y^{t})}\right]^{1/2} \tag{5}$$

(b) Efficiency Change (EC) is an efficiency comparison between the efficiency of the production frontier of the t period and that of the t+1 period. Eq. (6) denotes the EC. If EC > 1, it indicates improved efficiency, whereas EC < 1 indicates reduced efficiency.

$$\begin{aligned}EC&=\left[\frac{A^{t+1}(X^{t+1},Y^{t+1})}{A^{t}(X^{t},Y^{t})}\right]\;/\;\left[\frac{D^{t+1}(X^{t+1},Y^{t+1})}{D^{t}(X^{t},Y^{t})}\right]\\&=\left[\frac{D^{t}(X^{t},Y^{t})}{D^{t+1}(X^{t+1},Y^{t+1})}\right]\times\left[\frac{A^{t+1}(X^{t+1},Y^{t+1})}{A^{t}(X^{t},Y^{t})}\right]\end{aligned} \tag{6}$$

(c) The Malmquist Productivity Index (M) is denoted by Eq. (7), which is a product of Eq. (5) and Eq. (6). M > 1 indicates improved productivity, whereas M < 1 indicates decreased productivity.

$$\begin{aligned}M_{t,t+1}&=TC_{t,t+1}\times EC_{t,t+1}\\&=\left[\frac{D^{t+1}(X^{t},Y^{t})}{D^{t+1}(X^{t+1},Y^{t+1})}\;\frac{D^{t}(X^{t},Y^{t})}{D^{t}(X^{t+1},Y^{t+1})}\right]^{1/2}\times\left(\left[\frac{D^{t}(X^{t},Y^{t})}{D^{t+1}(X^{t+1},Y^{t+1})}\right]\times\left[\frac{A^{t+1}(X^{t+1},Y^{t+1})}{A^{t}(X^{t},Y^{t})}\right]\right)\\&=\left[\frac{D^{t+1}(X^{t},Y^{t})}{D^{t+1}(X^{t+1},Y^{t+1})}\;\frac{D^{t}(X^{t},Y^{t})}{D^{t}(X^{t+1},Y^{t+1})}\right]^{1/2}\times\left[\frac{A^{t+1}(X^{t+1},Y^{t+1})}{A^{t}(X^{t},Y^{t})}\right]\end{aligned} \tag{7}$$

4. Emperical Study: Taiwan's 38 Upgraded Technical Institutes

This section analyzes the managerial efficiencies of Taiwan's 38 technological institutes which were upgraded from junior college status between 1995 and 1998. The outcomes of this section includes two subsections: (1) relative efficiency analysis by current-period, shown in Table 2-3, and (2) performance variation analysis by cross-period, shown in Table 4.

4.1 Relative Efficiency Analysis by Current-period

The ANOVA results are as follows:

(1) According to the school type analysis result, the managerial efficiencies of private schools are superior to those of public schools, with the differences under a significant level at a = 0.05.

(2) There is no significant difference between managerial efficiencies of public and private schools in the analysis according to geographical location. This means that geographical location will not cause school managerial efficiency differentiation.

(3) According to the results of school size analysis, with a significant level at a = 0.05, the managerial efficiencies of larger schools, having 201 classes and above, are superior to those of smaller schools. This indicates that school managerial efficiency is relates to school size.

4.2 Cross-period Efficiency Analysis

Both technical change (*TC*) and efficiency change (*EC*), spanning 1995 and 1998, can be calculated through these four distance function values. The product of *TC* and *EC* is the Malmquist productivity index (*M*). If $M > 1$, it indicates improved efficiency, in other words, the productivity of specific institute increased over the previous four years; and if $M < 1$, it indicates reduced efficiency, in other words, the productivity of specific institute decreases over the previous four years.

The average value of the Malmquist productivity index for these institutes is 0.9855; indicating that the efficiency performance during the target period was reducing, which was caused primarily by the change of production frontier; while the technical efficiency change (*EC*) of these institutes during the target period was increasing. The average value of the technical change of the 38 institutes overall, was 0.9043, which is slightly declining; indicating the overall production technique of all sample institutes is recessing. The average value of the *EC* is 1.1111, which is increasing; indicating the overall efficiency variation has improved in comparison with the referenced period.

4.3 Analysis of Managerial Decision-making

This matrix can play the role of managerial decision-making matrix of further improvement efforts. Four groups of technological institutes are described below:

A. "Star group" includes Mingchi, Southern Taiwan, Chengshiu, Lien Ho, Chia Nan, Van Nung, and Ta-Hwa Institute of Technology. They are role models for the other technological institutes and have achieved outstanding managerial efficiency in the past and currently. They will be able to stay in the leading position if they control better those suggested resource indices and avoid committing vital administrative mistakes.

B. "Potential group" includes Fooyin, Huwei, Yuanpei, Tajen, Kung Shan, Chungtai, and Chien Hsin Institute of Technology. Their current resource input does not have efficient output, although their effort invested over the past four years produced positive production efficiency. It is possible for them to move up into the "star group" if they make improvements.

C. "General group" includes Dahan, Kao Yuan, China, Tainan Woman's, Lunghwa, Hungkuang, Far East, Chien Kuo, Mingshin, Yung Ta, Chung Hwa, Jin-Wen, Ling Tung, Wen Tzao Ursuline, Fortune, and Chin-Yi Institute of Technology. Their current resource input has not brought about efficient output, their efforts invested over the past four years have not achieved cause satisfactory efficiency, and they do not have the competitive advantages of the institutes of the "potential group". Suggested strategies would be: revising mid-range plans to develop unique development strengths and re-scheduling the application of resource input.

D. "Left-behind group" includes Ilan, Kaohsiung, Tzu Chi, Taichung, Chiayi, Pingtung, Kaohsiung, and St. John's & St. Mary's Institute of Technology. They currently have low relative efficiency in present. Recommended improvement approaches include: a complete revision to their short-, middle-, and long-term plans, as well as continued efforts to determine feasible endeavors.

5. Conclusions

Analyses incorporated the influences of school type, geographical location, and school size on school managerial efficiency performances. The analysis results indicated the following three main points: first, there is a significant correlation between school types and relative efficiency; seconds, there is a significant correlation between school size and relative efficiency; third, there is no significant correlation between geographical location and managerial efficiency, this indicates that the geographical location will not cause significant differences in managerial efficiency. According to the analysis results of cross-period efficiency variation, the technical changes one slightly increasing in Taiwan's technical institutes which have been upgraded from junior colleges. This means there is an improvement of overall technical level with a minimal scale. The improvement of efficiency change is higher than the technical change; indicating that there is a recession of school production technique. Matrices of this type can help site management to design more appropriate improvement strategies and to accomplish greater improvements. They can also help educational authorities officials to present a more appropriate evaluation system, which will provide accurate feedback information, making a professional school evaluation and guidance possible.

References

1. Ahn, T., Arnold, V., Charnes, A., and Cooper, W. W. (1989). "DEA and ratio efficiency analysis for public institutions of higher learning in Texas." Research in Governmental and Nonprofit Accounting, 5(2), 65-185.
2. Banker, R. D., Conrad, R. F., and Strauss, R. P. (1986). "A comparative application of data envelopment analysis and translog method: an illustrative study of hospital production," Management Science, 32(1), 30-44.
3. Bates, J. M., (1997). "Measuring predetermined socioeconomic input when assessing the efficiency of educational output," Applied Economics, 29(1), 85-93.

4. Bessent, A., Bessent W., Charnes, A., Cooper, W. W., and Thorogood, N. (1983). "Evaluation of educational program proposals by means of data envelopment analysis," Educational Administration Quarterly, 19(2), 82-107.
5. Caves, D., L. Christensen, and E. Diewert. (1982). "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity", Econometrica 50, 1393-1414.
6. Chang, Tien-Fu (1995). Study for the Construction of Educational Index in Taiwan Area, Taipei, national Science Council.
7. Charnes, A., Cooper, W. W., and Rhodes, E., (1978). "Measuring the efficiency of decision making units," European Journal of Operational Research, 2(6), 429-444.
8. Charnes A., Cooper W. W., and Rhode E., (1981). "Evaluating Program and Managerial Efficiency: An application of DEA to program follow through," Management Science, 27(6) 668-697.
9. Fare, R., Grosskopf, S., Lindgren, B., and Roos P., (1992). "Productivity Changes in Swedish Pharmacies 1980-1989: A Non-Parametric Malmquist Approach, " The Journal of Productivity Analysis, 3(1), 85-101.
10. Higgins, J. C. (1989). "Performance measurement in universities," European Journal of Operational Research, 38, 358-368.
11.Hufner, K. (1987). "Differentiation and competition in higher education: recent trends in the Federal Republic of Germany." European Journal of Education, 22(2), 133-143.
12.Inbar, D. E. (1988). "Quality educational indicators in a nation in the making: The case of Israel." Studies in Educational Evaluation, 14(1), 55-63.
13.Kao, C., (1994). "Evaluation of junior colleges of technology : The Taiwan case," European Journal of Operational Research, 72(1), 43-51.
14.Miguel, M. M., and Ordenez, V. M. (1988). "The Philippine elementary school system." Studies in Educational Evaluation, 14(1), 37-45.
15.Power, C. (1989). "Assessing the effectiveness of secondary schooling in Australia." Studies in Educational Evaluation, 15(1), 47-71.
16.Wang, Pao-Chin (1993). The Study of the Performance Index of Higher Education. Ph.D. Dissertation, Graduate School of Education, ChengChi University.
17.Wang, Pao-Chin (1995). The Study of Feasible Model for Assessment Performance Index of Higher Education, Publication of National Science Council: Humanity and Social Science, 6(1), 127 – 147.

Using DEA of REM and EAM for Efficiency Assessment of Technology Institutes Upgraded from Junior Colleges: The Case in Taiwan

Li-Chen Liu, Chuan Lee, and Gwo-Hshiung Tzeng*

Graduate Institute of Management Science, Ming Chuan University, Taipei, Taiwan; *Institute of Management of Technology, National Chiao Tung University, Hsinchu, Taiwan; e-mail: ghtzeng@cc.nctu.edu.tw

Abstract. *The purpose of this study is to improve the assessment system for vocational education, thus helping to increase technology college education efficiency. This study employs both classical and new data envelopment analysis (DEA) to study relative radial efficiency, management efficiency and scale efficiency of the application of resources for the 38 technology colleges in Taiwan during 1998. Analysis outcomes showed that ranking differentiation between relative classical radial efficiency measure (REM) and efficiency achievement measure (EAM) is consistent, if the same group index differentiation, of an institution, is consistent. Both models support each other and provide clearer relative Pareto efficiency and efficiency achievement for comparison of results.*
Keywords: data envelopment analysis, efficiency achievement measure, management efficiency, scale efficiency, multiple objective programming, technology college

1. Introduction

Currently, Data Envelopment Analysis (DEA) is an assessment approach in management science being extensively adopted to assess the efficiency of multiple-input and multiple-output systems. There are many successful cases in non-profit businesses such as schools, libraries, and public hospitals in addition to for-profit business such as banks and hotels. However, it has not been widely used for technology colleges.

According to the data collected, both classical and new DEA approaches (Charnes et al., 1978; Chiang and Tzeng, 2000) are used to explore the management efficiency of these technology colleges. The results indicate that the efficiency measure approach of classical DEA belongs to relative Pareto's efficiency measure approach. Along with efficiency achievement measure, this can further provide more accurate results for relative Pareto efficiency and efficiency achievement for comparisons. It can also improve the measurement of technology college management efficiency, which is useful for policy-making in these institutions.

There are five sections following; the second section illustrates how sample and variables were selected for efficiency assessment. The third section illustrates both the assessment of classical DEA.

2. Selection of Variables and Samples for Efficiency Assessment

Both classical DEA and new Fuzzy multiple objective DEA were employed to explore the organizational goal and selected input and output relation of the 38 junior colleges that reformed and upgraded to technology colleges in Taiwan.

Since otherwise researchers' ideology and interests, as well as the models employed could create variation in the definition of input and output. Previously, Hufner (1987) listed indexes to assess the reputation of universities, including the number of full-time teachers, teacher-student ratio, number of passed doctoral dissertation published, number of published papers for qualifying for the position of professor, and number of trusted research cases. Inbar (1988) listed percentage of GNP taken by educational expanses and percentage of public expenses taken by education expenses as input indexes, with the number of graduate, mathematic competence, and language ability as three output indexes.

The indexes selected in this study cover those indexes mentioned above, and it also considers both the feasibility of practical operation and the possibility of obtaining data. Educational input variables covered four phases that are physical, human resource, facility, and information. Input variables are the total floor area of school buildings; number of faculty above, including assistant professor; expenditure; volumes of books and categories of journals in the library. Outputs variables are divided into the three phases of teaching efficiency, practical research efficiency, and service efficiency. Output assessment indexes are numbers of graduates, total amount of faculty's research fund, and income from the co-op program and extension education.

3. Measure of Assessment Model

This study adopted both the classical Data Envelopment Analysis proposed by Charnes et al. (1978) and the new DEA Fuzzy Multiple Objective Programming Approach proposed by Chiang and Tzeng (2000) to conduct assessment of relative and absolute rank.

3.1 Classical Radial Efficiency Measure

According to the CCR Model proposed by Charnes, et al. in 1978:

$$\text{Max } h_a = \frac{\sum_{r=1}^{S} u_r y_{ra}}{\sum_{i=1}^{m} v_i x_{ia}} \tag{1}$$

subject to:

$$\frac{\sum_{r=1}^{S} u_r y_{rk}}{\sum_{i=1`}^{m} v_i x_{ik}} \leq 1, \quad k = 1,2,\ldots,n$$

$$0 < e \leq u_r, 0 < e \leq v_i, \quad i = 1,\ldots,m, \quad r = 1,\ldots, s$$

where x_{ik} stands for the i^{th} input of the k^{th} DMU; y_{rk} stands for the r^{th} output of k^{th} DMU; u_r, v_i stand for the weight of the r^{th} output and i^{th} input respectively; h_a is relative efficiency value. Since Eq.(1) involves fractional programming, that is difficult to solve, Charnes et al. (1978) therefore, converted it to linear programming (LP) model in order to find solution.

$$\text{Max } h_a = \sum_{i=1}^{S} u_r y_{ra} \tag{2}$$

subject to:

$$\sum_{i=1}^{m} v_i x_{ik} - \sum_{r=1}^{s} u_r y_{rk} \geq 0, \quad k = 1,\ldots,n$$

$$\sum_{i=1}^{m} v_i x_{ia} = 1$$

$$u_r \geq \varepsilon > 0, \quad r = 1,\ldots,s, \qquad v_i \geq \varepsilon > 0, \quad i = 1,\ldots,m$$

The number of elements in the dual problem in Eq. (2) can be reduced for the sake of finding answer, the dual problem after conversion becomes:

$$\text{Min } \{?_a - e[\sum_{i=1}^{m} S_{ia}^- + \sum_{r=1}^{s} S_{ra}^+]\} \tag{3}$$

subject to:

$$\theta_a x_{ia} - \sum_{k=1}^{n} \lambda_k x_{ik} - S_{ia}^- = 0, \quad i = 1,\ldots,m$$

$$y_{ra} - \sum_{k=1}^{n} \lambda_k y_{rk} + S_{ra}^+ = 0, \quad r = 1,\ldots, s$$

$$S_{ia}^-, S_{ra}^+, \lambda_k \geq 0$$

where S_{ia}^- and S_{ra}^+ are slack variables.

3.2 Multiple Objective Programming Assessment Model for Efficiency Reach

To achieve this goal, the concept of multiple objective programming can be employed to find a set of consistent weight combination approach so that the optimized efficiency value can be calculated for each DMU_k in overall relative efficiency achievement. This idea can be formulated as Eq.(4).

$$\text{Max}\left[h_1 = \frac{\sum_{r=1}^{s} u_r y_{r1}}{\sum_{i=1}^{m} v_i x_{i1}}, h_2 = \frac{\sum_{r=1}^{s} u_r y_{r2}}{\sum_{i=1}^{m} v_i x_{i2}}, \ldots, h_n = \frac{\sum_{r=1}^{s} u_r y_{rn}}{\sum_{i=1}^{m} v_i x_{in}} \right] \tag{4}$$

subject to:

$$\frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} \leq 1 \ , \ k = 1,2,\ldots,n$$

$$u_r \geq \varepsilon > 0, \ r = 1,2,\ldots,s, \quad v_i \geq \varepsilon > 0, \ i = 1,2,\ldots,m$$

Given a linear identity function, h_k^L and h_k^R, donate the left and right frontier value of h_k, the k^{th} objective function value, respectively. The span of h_k^L and h_k^R lies between 0 and 1 due to the outcome of the objective equation of Eq.(4) is efficiency ratio, that is $h_k^L = 0$ and $h_k^R = 1$, and the identity value is $u(h_k)$, which can be considered as the achieved value of efficiency ratio h_k for the DMU. No doubt that the value is between 0 and 1. Such a function is called an identity function. Therefore, Eq. (4) can be converted to the pattern of fuzzy multiple objective programming (Tzeng and Chen, 1999; Chen and Tzeng, 1999)

$$\text{Max} \ \alpha \tag{5}$$

subject to:

$$\sum_{r=1}^{s} u_r y_{rk} - \sum_{i=1}^{m} v_i x_{ik} \leq 0, \ k = 1,\ldots,n$$

$$\sum_{r=1}^{s} u_r \cdot y_{rk} - a \cdot \sum_{i=1}^{m} v_i \cdot x_{ik} \geq 0, \ k = 1,\ldots,n$$

$$0 < a \leq 1; \quad u_r \geq e > 0, \ r = 1,\ldots,s, \quad v_i \geq e > 0, \ i = 1,\ldots,m$$

A set of (u^*,v^*) can be calculated according to Eq. (5) and the efficiency value h_k of a DMU can be calculated with the value of (u^*,v^*). Since the efficiency value h_k of each DMU actually equals its efficiency achievement of the efficiency value h_k (because satisfaction is assumed to be identity function), therefore, we can define the efficiency achievement measure as:

$$\alpha_k = \sum_{r=1}^{s} u_r^* \cdot y_{rk} / \sum_{i=1}^{m} v_i^* \cdot x_{ik} \tag{6}$$

4. Analysis and Conclusion for the Results of Case Study

This section is divided into three subsection:

(1) Radial Efficiency Measure

At a =0.05, private schools perform better than public schools, in terms of school management efficiency; the difference of the management efficiency of schools, at different geographic location is insignificant; and the management efficiencies of large-sized schools, more than 201 classes, are better.

(2) Measurement of Scale Efficiency

There are two technology colleges, Chin-Yi and Lien Ho, that achieved Pareto optimal organization, although this does not mean there is no room for them to improve. Because DEA is a concept of relative comparison, therefore, if one of the two colleges has improvement in some area, the assessment grade could be less than 1 and become Pareto non-optimal organization.

Among those technology colleges that has less than 100 classes, there are two technology colleges, Mingchi and Wen Tzao, that achieved POO and both were ranked as No. 1; the other four are PNO and are ranked by aggregate efficiency. Six technology colleges are in the stage of "Increasing Return to Scale". Decision makers of these schools may try to improve operational efficiency through expanding school size.

(3) Efficiency Achievement Measure and Comparative Analysis with Radial Efficiency

Applying the efficiency achievement multiple objective programming model, only three schools turn out to be relatively efficient, among then Chia Nan ($\alpha_k = 1$), Lien Ho ($\alpha_k = 1$), and Van Nung Institute of Technology ($\alpha_k = 1$), the rest part of subject schools are relatively inefficient, such an approach is similar with the De Novo programming approach (Zeleny, 1986, 1995), which is capable of breaking through the barriers of Pareto's solution and provid more room for further development.

Furthermore, this study applied cluster analysis after normalizing the information to categorize the 38 technology colleges into seven groups by their characteristics. The effectiveness of efficiency ranking obtained from both efficiency achievement measure and radial efficiency measure are quite similar.

By efficiency achievement measure, we can further find that nine technology colleges including Southern Taiwan, Chengshiu, Mingchi, Wen Tzao, Ta-Hwa, and Hungkuang were relatively inefficient compared to Chia Nan College of Pharmacy and Science, Lien Ho, and Van Nung. This provides more accurate results for comparing the relative efficiency ratio and absolute achievement ratio.

6.Conclusions

The results of case analysis shows that the efficiency measure approach of classical DEA is in the relative Pareto's efficiency measure approach. The results of this show that 12 institutes of technology are relatively efficient; that is, the result of radial efficiency measure are 1. Along with efficiency achievement measure, this can further provide more accurate results for relative Pareto efficiency and efficiency achievement for comparisons.

References

Charnes, A., Cooper, W. W., and Rhodes, E., (1978). "Measuring the efficiency of decision making units," European Journal of Operational Research, 2(6), 429-444.

Chen, Y. W. and Tzeng, G. H. (1999). "A fuzzy multi-objective model for reconstructing post-earthquake road-network by genetic algorithm," International Journal of Fuzzy Systems, 1(2), 45-55.

Chiang, C. I. and Tzeng, G. H. (2000). "A New Efficiency Measure for DEA : Efficiency Achievement Measure Established on Fuzzy Multiple Objectives Programming," Journal of Management, 17(2), 369-388.

Chiang, C. I. and Tzeng, G. H. (2000). A multiple objective programming approach to data envelopment analysis, Shi, Y. and Zeleny, M. (eds) "New Frontiers of Decision Making for the Information Technology ERA," World Scientific Publishing Company, 270-285.

Hufner, K. (1987). Differentiation and competition in higher education: recent trends in the Federal Republic of Germany. European Journal of Education, 22(2), 133-143.

Inbar, D. E. (1988). Quality educational indicators in a nation in the making: The case of Israel. Studies in Educational Evaluation, 14(1), 55-63.

Tzeng, G. H. and Chen, Y. W. (1999). "The optimal location of airport forestations: a fuzzy multi-objective programming through revised genetic algorithm," Transportation Planning and Technology, 23(1), 37-55.

The Comprehensive Financial Risk Management in a Bank – Stochastic Goal Programming Optimization*

Jerzy Michnik

The Karol Adamiecki University of Economics in Katowice
ul. Bogucicka 14, 40-227 Katowice, Poland
e-mail:jmichnik@ae.katowice.pl

Abstract. Risk management have developed substantially so it became a distinctive subfield in the theory of finance. As the banks operate on the large scale on the financial markets, the risk management is an important issue in the banking industry [1].

This paper presents the stochastic extension of the deterministic model originally developed by the author. The model comprises main financial risks and allows the use of most important financial instruments, so it can be regarded as a basic module of the risk management support system in a commercial bank.

1 Introduction

There is a long history of optimization models designed as a support for decision making in the field of financial risk management in the commercial banks. Usually the models were limited to one type of risk and/or single period [2]-[4]. Some projects attained practical realizations. The examples are: linear goal programming model of Giokas and Vassiloglou [5] and two-stage linear goal programming model of Korhonen [8]. Both these models were based on balance sheet data so they were designed rather to support financial planning then to risk management itself.

The deterministic multicriterial model especially designed for financial risk management was developed by the author [9,10]. The model was based on future cash-flows and comprised the main financial risks and allowed the use of wide variety of financial instruments. The stochastic, single-objective linear programming model, founded on the same framework, was presented by the author last year [11]. This work presents the union of these previous ideas – the model was developed to be multiobjective and to has stochastic nature as well.

* Research partly supported from a KBN Grant no. 2 H02B 006 22 "The Multicriteria Methods on Polish Financial Market.

2 Model formulation

We assume that the risk the bank is exposed to, is a result of the uncertainty of the time structure of future cash-flows. Let's consider the finite period of time. We divide it into finite number of periods and assume that all cash-flows take place only at the end of any period. Then every instrument is characterized by the set of its cash-flows $\breve{x}_t(v,z)$, i.e.

$$\breve{x} = [\breve{x}_t(v,z)]_{t\in\mathcal{T}} \tag{1}$$

where $\mathcal{T} = \{0, 1, \ldots, t_n\}$, $v \in \mathcal{V}$ $z \in \mathcal{Z}$. The index t is used for numbering the ends of periods, v for currency, while z refers to all other characteristics of given instrument. The randomness come into the model through the two sets of parameters: interest rates – $\tilde{r}_t(v,z)$ and exchange rates – $\tilde{q}_t(v_1, v_2)$ (currency v_1 against currency v_2).

2.1 Decision variables

As the decision variable connected with given instrument, we will use single non-negative number $x_t(v,z)$, from which the all cash-flows can be derived. For example, in case of purchase of interest instrument with coupons, decision variable $x_t(v,z)$ is a volume which is purchased. It generates $-x_t(v,z)$ flow at a moment of purchase and several interest flows of the form $\tilde{r}_{t'}(v,z)x_t(v,z)$ in later moments t', and the return of the capital at maturity. When we consider forward contract or derivative instrument the future cash-flows have an optional form. Lets consider purchase of interest rate option as an example. The principal amount of the option x_t $(t > 0)$ is the decision variable (for simplicity we omit for the moment v and z). The initial cash-flow is $-r_0 x_t$, where r_0 is the unit price of the option (expressed as percentage) The conditional flow in the future, which depends on the actual value of the interest rate at moment t, will be

$$x_t(\tilde{r}_t - \hat{r}_t)^+, \tag{2}$$

where $\hat{r}$ is exercise rate of the option[1].

2.2 Constraints

The external and internal economic conditions, which a bank operates in, impose several constraints of different nature and structure. We grouped them

[1] We use following notation:

$$a^+ = \begin{cases} a & if \quad a \geq 0 \\ 0 & otherwise \end{cases}$$

according to their essential meaning. Market and technical limits arise from the situation of a bank in it's economical environment. These limits has to be proposed by the bank specialists and confirmed by the management of a bank. Some of them are simply top limits for bank's dealers transactions and open positions. For this limits we assume the form

$$Bx \leq b, \tag{3}$$

where matrix B and vector b are deterministic.

The other important group contains all the constraints which are imposed by the legal system of the country, the bank operates in. The detailed rules are usually issued by banking supervisory. To this group belong e.g.: the bottom limits for the capital adequacy ratio and for cash reserves. The latter ones can be regarded as the part of risk management constraints as well, because it is common that the management of a bank impose more strict conditions on cash reserves, dependent on bank's situation. The limit for capital adequacy ratio has the form (3), while the limits for cash reserves has the stochastic character. This type of constraints can be written as

$$H(\tilde{r}, \tilde{q})x \leq h(\tilde{r}, \tilde{q}) \tag{4}$$

where both matrix H and vector h generally depend on stochastic parameters $\tilde{r}$ and $\tilde{q}$.

2.3 Criteria

When the problem of risk management is of concern, the decision maker are to minimize the several types of risk and achieve the maximum possible profit at the same time. This guidelines the contents of typical optimization model. In the financial management of the commercial bank, the interest rate and foreign exchange risk are of most interest. The popular method of controlling the interest risk is the duration gap management. We take as the criterion the minimization of risk of market value of bank's capital. This is the only criterion based on duration which is linear form of decision variables. Let's define

$$D_X^n = D_X PV_X = \sum_t \gamma_t \, cf_{X,t}(x), \tag{5}$$

where $X = A$ for assets, $X = L$ for liabilities,
D_X,PV_X and $cf_{X,t}(x)$ - duration, present value and sum of all positive (negative) cash-flows in moment t of assets (liabilities), respectively, γ_t – discount factor. As the risk is minimal for zero gap, in the goal programming frame we receive the constraint:

$$D_L^n - D_A^n - y_r^+ + y_r^- = 0, \tag{6}$$

where y_r^+, y_r^- are over- and underachievement auxiliary variables.

The foreign exchange risk with can be controlled via foreign currency position. It can be defined it as follows:

$$P_t(v) = p_t(v) + \sum_{z \in \mathcal{Z}_f} x_t(v, z), \quad t \in \mathcal{T},\ v \in \mathcal{V}', \tag{7}$$

where $P_t(v)$ – position for period t and for foreign currency v, $p_t(v)$ – initial position for period t and currency v, $\mathcal{Z} \supset \mathcal{Z}_f$ – transactions changing position in a given foreign currency, $\mathcal{V}' = \mathcal{V} \setminus \{v_0\}$, v_0 – local currency.

Again the minimal risk is when the position is zero (closed), so we can formulate another set of linear constraints for goal programming model

$$P_t(v) - y_{f,t}^+(v) + y_{f,t}^-(v) = 0, \quad t \in \mathcal{T},\ v \in \mathcal{V}'. \tag{8}$$

As it was said above, we need some profitability criterion as a counterpart of risk minimization criteria. The profit for interest instrument can be written as simple as $\tilde{r}_t(v,z)x_{t'}(v,z)$. For derivative instruments the profit has the form $x_t[(\tilde{r}_t - \hat{r}_t)^+ - r_0]$. We use general form $R[\tilde{r}_t(v,z)]$ for these expressions. If we assume, that $R[\tilde{r}_t(v,z)] = 1$ for foreign exchange transactions, we will be able to write down every contribution to profit as single expression $R[\tilde{r}_t(v,z)]x_{t'}(v,z)$. Now, all positive and negative contributions to profit has to be translated to single currency (usually local one), discounted to one point in time and added together in the end. The resulting constraint for goal programming model is

$$\sum_{v \in \mathcal{V}} \sum_{z \in \mathcal{Z}} \sum_{t,t' \in \mathcal{T}} \gamma_t\, R[\tilde{r}_t(v,z)]\, x_{t'}(v,z)\, \tilde{q}_t(v,v_0) - y_p^+ + y_p^- = P, \tag{9}$$

where P is the desirable level of profit.

2.4 Stochastic model and it's deterministic equivalent

The presented model takes the following form

$$\begin{aligned} \min\nolimits_{x,y}\, [c^+y^+ + c^-y^-] & \\ \text{s.t.} & \\ Bx \le b & \\ H(\tilde{r},\tilde{q})x \le h(\tilde{r},\tilde{q}) & \\ D(\tilde{r},\tilde{q})x - y^+ + y^- = G & \end{aligned} \tag{10}$$

This is the stochastic linear goal programming model and such a model can not be solved explicitly (for general reference see the monography [7]). For solving it we need to transform it to deterministic equivalent form. It was shown by Heras and Aguado [6], that stochastic goal programming model is a particular case of stochastic linear programming with recourse. In our case it is convenient to formulate it as the *multistage recourse program.* We need additional auxiliary variables y' which serve to compensate the violation of stochastic constraints in some realizations of random parameters. The number of variables y' is equal the number of stochastic inequalities in (10).

The deterministic equivalent of our model is

$$\begin{aligned} \min_x E_{\tilde{r},\tilde{q}}\{Q_1(x,\tilde{r},\tilde{q}) + Q_2(x,y',\tilde{r},\tilde{q})\} \\ \text{s.t.} \\ Bx \leq b \\ H(\tilde{r},\tilde{q})x - y' \leq h(\tilde{r},\tilde{q}) \\ D(\tilde{r},\tilde{q})x - y^+ + y^- = G \end{aligned} \tag{11}$$

where $E_{\tilde{r},\tilde{q}}$ – expectation with respect to the distributions of $\tilde{r}$, $\tilde{q}$.

$Q_2(x,y',\tilde{r},\tilde{q})$ in (11) is the third stage recourse function and is given by

$$Q_2(x,y',\tilde{r},\tilde{q}) = \min_y\{c^T y \mid y \geq 0\}, \tag{12}$$

The variables y in above equation are the usual over- and under-achievement variables in goal programming, consequently the coefficients c reflect the relative importance of different criteria for the decision maker.

$Q_1(x,\tilde{r},\tilde{q})$ in (11) is the second stage recourse function and is given by

$$Q_1(x,\tilde{r},\tilde{q}) = \min_{y'}\{d^T y' \mid y' \geq [H(r,q)x - h(r,q)]^+\} \tag{13}$$

Variables y' can be interpreted as the amounts of additional hedging transactions which have to be done whenever the particular stochastic constraint is violated as the result of concrete realization of random variables. Consequently coefficients d can be interpreted as interest or exchange rates of these hedging transactions and they are usually less favourable for the bank.

3 The exemplary model and computational tests

As there is no enough room for detailed description of the model we present here only brief picture of it[2].

There are a wide variety of financial tools commonly used in a financial risk management process. For the purpose of testing we use some typical examples which represent different types of instruments which have similar cash-flow structure. For every instrument we have in the model a decision variable which generates all it's cash-flows.

Fixed rate spot transactions represent the purchase and sell of treasury bills and transactions with other banks on money market. The treasury bonds with coupon payments represent variable rate instruments. The other specific group are foreign exchange spot and forward transactions. Derivative

[2] Those who are interested in some details are requested to contact the author.

instruments contribute to the model through the call and put interest rate options.

The set of constraints comprised the top limits for every instrument, the bottom limits for capital adequacy ratio and cash reserves. Three criteria presented earlier, were incorporated: minimization of interest rate risk and foreign exchange risk and maximization of profitability.

For numerical tests there were chosen three periods ($t_n = 2$) and two currencies ($v = 2$), so the size of the model was reasonable. It had 30 decision variables and 30 auxiliary variables (y and y'). Calculations were made for different selection of coefficients c to observe different solutions for different decision maker attitudes. It proved the possibility to use this method as quasi-interactive technique which enables elastic reaction to decision maker needs.

4 Conclusions

The financial risk management in a commercial bank is a very complicated process which obeys a great number of elements - instruments and parameters. So it is clear that the any method of optimization type can be useful auxiliary tool in such process.

The proposal presented in this paper has the advantage of taking the random nature of main parameters which appear in the process and comprises a number of instruments used for risk control in commercial banks. It is quite general as it allows the wide choice of time horizon and its division into time periods – from days and weeks up to months or even years. So it can be easily adjusted for daily risk control operations, medium-time risk management or long-time strategic plans.

Multicriterial formulation of optimization models proved their dominance over single criterial formulation for a long time. Goal programming version gives the opportunity to use simple method for model solution.

References

1. Bessis, J. (1998): Risk Management in Banking. J. Wiley & Sons, Chichester
2. Bessler, W., Booth, G.G. (1994): An interest rate risk management model for commercial banks. European Journal of Operational Research 74, 243-256
3. Booth G.G., Bessler W., Foote W.G. (1989): Managing interest-rate risk in banking institutions. European Journal of Operational Research 41 (1989), 302-313
4. Booth G.G., Dash G.H., Jr. (1979): Alternate programming structures for bank portfolios. Journal of Banking and Finance 3 (1979), 67-82
5. Giokas, D., Vassiloglou, M. (1991): A goal programming model for bank assets and liabilities managament. European Journal of Operational Research 50, 48-60

6. Heras A., Aguado A.G. (1999). Stochastic Goal Programming. Central European Journal of Operations Research, vol. 7 no. 3, 139-158
7. Kall, P., Wallace, S.W. (1994): Stochastic Programming. J. Wiley & Sons, Chichester
8. Korhonen, A. (1987): A dynamic bank portfolio planning model with multiple scenarios, multiple goals and changing priorities. European Journal of Operational Research 30, 13-23
9. Michnik, J. (1998):A Multicriterial Model for Risk Management in a Commercial Bank. In: Trzaskalik, T.(Ed.): Preference Modeling and Risk. The Karol Adamiecki University of Economics in Katowice, 275-288 (in Polish)
10. Michnik, J. (1999): An Application of Interactive Goal Programming in Risk Management in a Bank. In: Trzaskalik, T.(Ed.): Preference Modeling and Risk. The Karol Adamiecki University of Economics in Katowice, 271-290 (in Polish)
11. Michnik J. (2001). The Stochastic Optimization of Financial Management in a Bank, lecture presented at Symposium On Operations Research 2001, Duisburg, Germany, 3-5.09.2001
12. Uyemura D.G., Van Deventer D.R. (1993): Financial Risk Management in Banking. Irwin, Chicago

The Effectiveness of the Balanced Scorecard Framework for E-Commerce

Jamshed J. Mistry, B. K. Pathak

Department of Management, Worcester Polytechnic Institute, Worcester, MA 01609 USA

Abstract: This paper reports evidence on the effectiveness of the balanced scorecard framework (BSC) in measuring and monitoring the performance of e-commerce companies. The study utilizes an integrated Data Envelopment Analysis (DEA) model to examine and evaluate the relative efficiency of the measures identified within the BSC framework for measuring the performance of E-Commerce companies. Finally, the study examines the effectiveness of the BSC framework in predicting the success or failure of E-commerce companies.

Keywords: Data Envelopment Analysis (DEA); Balanced Scorecard; Performance Measurement; Electronic Commerce.

1. Introduction

In the last decade, researchers have become increasingly interested in analyzing the impact of non-financial performance measures on firm performance. Kaplan and Norton (1992) proposed the Balanced Scorecard framework, which emphasizes the need to measure both financial and non-financial parameters of performance. Non-financial measures are particularly relevant for new age companies (dot.coms, business to consumer companies) that have revolutionized the marketplace and appear to defy the basic rules of business. Dot-com or E-commerce companies have necessitated the development of a whole new set of performance measurement parameters for monitoring and measuring their performance (Seybold & Marshak, 1999).

Although the Balanced Scorecard model was initially proposed in 1992, and the model has been widely accepted by most practitioners, little empirical analysis has focused on validating the model. In this paper, we develop four sets of performance measurement parameters specifically designed for E-commerce companies, drawing on the Balanced Scorecard framework. We then employ Data Envelopment Analysis (DEA) using these measures to examine the efficiency of Balanced Scorecard parameters in measuring the performance of eighteen E-commerce companies. Finally, we focus on six of the eighteen companies to compare the three most successful companies with three that subsequently failed in order to examine the effectiveness of the Balanced Scorecard parameters in predicting bankruptcy.

2. Background and Significance

Kaplan and Norton (1992) used a balanced scorecard, which requires managers to balance four different but linked perspectives in order to identify appropriate measures of performance. The first perspective represents traditional accounting measures that report the financial consequences of actions already taken. This financial perspective highlights how the company appears to shareholders and concentrates on measures relating to profitability and growth, cash flow and gearing. The Balanced Scorecard supplements these financial measures with three other perspectives dealing with (a) customers, (b) internal processes, and (c) the firm's innovation and learning record - all three areas that are important drivers of future financial performance. The customer perspective is designed to highlight the factors that really matter to customers such as value for money, time and performance. The internal business perspective is designed to focus on those critical business activities that must be performed in order to satisfy the expectations of its customers. These include cycle time, quality and efficiency of operations. The innovation and learning perspective highlights the fact that, in the face of intense competition, firms must make continual improvement and have the ability to introduce new products in the future.

Measuring the performance of E-commerce companies has always been a relatively difficult task. E-commerce firms have focused on innovation continuously in order to integrate technology with offering customized tailor-made services to the customers and use parameters such as revenue, click through ratios and other indirect parameters to measure their performance. Thus E commerce companies measure their performance by using a mix of traditional and new parameters. For example, McKinsey's e-performance scorecard, launched in 1999, collects data about a variety of visitor, customer, and financial metrics (Agrawal, Arjona & Lemmens, 2001). The scorecard comprises 21 indicators that measure performance both statically (at one point in time) and dynamically (over a period of time). These indicators are grouped into three categories—attraction, conversion, and retention—and then folded into the overall e-performance scorecard, which is a weighted average of the twenty-one indicators. The scorecard highlights two key dimensions: the efficiency of costs (for example, the cost of attracting visitors to a site and of maintaining active customers) and the effectiveness of a site's operations (such as conversion rates, the rate at which the number of customers increases, and customer gross margins). Best practice in the e-business sector combines the lowest costs with the highest effectiveness.

Thus, parameters such as the customer point of view and integration of technology to produce personalized web content for customers are important measures of performance in E-commerce companies. Therefore, we focus on both financial and non-financial measures of performance in E-commerce companies. We utilize DEA analysis to examine the efficiency of eighteen such companies on each of the BSC parameters.

3. Methodology

Measures of Balanced Scorecard dimensions: For each dimension of the BSC, we identified performance indicators developed for E-commerce companies, and selected specific measures based on the research mentioned in the previous section. The specific measures are presented in Table 1.

Table 1. Measures for the BSC dimensions

Perspective	Inputs		Outputs		
Customer	Marketing Expenditure		Revenue	Number of Customers	Number of visitors
Internal Processes	Number of Employees	Financing	Revenue	Number of Customers	Number of visitors
Inn./ Know. Mgmt.	Number of Employees	Tech./ Dev. Expenditure	Revenue	Number of Customers	
Finance	Financing	Net Income		Revenue	

Sample: The data were obtained from the E-commerce Almanac data set collected by the Intermarket Group. This almanac compiles exhaustive information about E-commerce companies and includes financial, marketing, operational and other information that can be categorized into the balanced scorecard framework. The original data set included eighty-two E-commerce companies. However, data on all the performance measures we derived for each of the four BSC dimensions were available for only eighteen companies.

Analyses: In the first set of analyses, DEA methodology was utilized to examine the efficiency of the eighteen companies on each of the four dimensions. Through the optimization for each individual unit, DEA yields an efficient frontier that represents and estimates the relations among the multiple performance measures (Charnes, Cooper and Rhodes, 1978).

Suppose we have a set of n decision making units (DMUs) (e.g., companies), DMU_j (j = 1, ..., n) and let x_i (i = 1, ..., m) be the m input performance measures where smaller values are preferred, e.g., cost measures and y_r (r = 1, ..., s) be the s output performance measures where larger values are preferred, e.g., revenue. Thus, we have m+s performance measures for the n DMUs. Further, we have x_{ij} as the observed value on the ith input performance measure and y_{rj} as the observed value on the rth output performance measure.

Based upon the observations, we have the following DEA model for evaluating the relative efficiency of DMU_o among other DMUs

$$
\begin{aligned}
& \min \theta \\
s.t.\quad & \sum_{j=1}^{m} \lambda_j x_{ij} \leq \theta x_{io} \quad i = 1,\ldots,m \\
& \sum_{j=1}^{s} \lambda_j y_{rj} \geq y_{ro} \quad r = 1,\ldots,s \qquad (1) \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0, \quad j = 1,\ldots,n
\end{aligned}
$$

Model (1) is called variable returns to scale (VRS) model in DEA (Banker, Charnes and Cooper, 1984). Model (1) is input-oriented, since it minimizes inputs while keeping the outputs at their current levels. We have an output-oriented model (for the negative input values in the finance perspective), which maximizes outputs while keeping the inputs at their current levels.

$$
\begin{aligned}
& \max \phi \\
s.t.\quad & \sum_{j=1}^{m} \lambda_j x_{ij} \leq x_{io} \quad i = 1,\ldots,m \\
& \sum_{j=1}^{s} \lambda_j y_{rj} \geq \phi y_{ro} \quad r = 1,\ldots,s \qquad (2) \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0, \quad j = 1,\ldots,n
\end{aligned}
$$

The above two models allow us to deal with negative inputs and outputs. See Zhu (2002) for additional DEA models.

4. Results

In our first set of analyses, we utilize DEA methodology to assess the efficiency of eighteen E-commerce companies on each of the four dimensions or perspectives of the BSC. While the finance perspective utilizes traditional financial measures, the customer, internal processes and innovation dimensions utilize non-financial measures. On each dimension, the companies are rank-ordered based on their efficiency scores for each dimension (See Table 2).

Table 2. Customer Perspective & Finance Perspective

Customer Perspective		Finance Perspective	
DMU Name	VRS Efficiency	DMU Name	VRS Efficiency
Amazon.com	1.00000	Amazon.com	1.00000
eBay	1.00000	eBay	1.00000
Buy.com	1.00000	E*Trade	1.00000
iPrint	1.00000	iPrint	1.03084
Peapod	1.00000	Peapod	1.05091
Priceline.com	0.93173	Outpost.com	1.06077
1-800-Flowers	0.82261	Furniture.com	1.07603
Webvan	0.61021	iOwn	1.08086
Nextcard	0.53494	Petsmart.com	1.08399
Outpost.com	0.53112	1-800-Flowers	1.08678
Cdnow	0.48576	CarsDirect.com	1.11428
CarsDirect.com	0.38828	Nextcard	1.12178
iOwn	0.37709	PlanetRX.com	1.15505
Beyond.com	0.31434	Buy.com	1.15614
E*Trade	0.26294	Cdnow	1.19104
Furniture.com	0.22682	Beyond.com	1.20081
Petsmart.com	0.22603	Priceline.com	1.20808
PlanetRX.com	0.13916	Webvan	1.22904

Since our objective was to examine the utility of the DEA efficiency scores in predicting future success or failure of the companies, we identified three companies that subsequently failed and three that remained successful and locate these six companies on the rank-ordered list in Table 2. Two successful companies (Amazon.com and ebay) and one of the ones that subsequently failed (Furniture.com), emerge as financially efficient in 1999, falling among the seven companies with the highest rank-orders (1-7). The two companies that subsequently failed (Webvan and PlanetRX.com) and one of the successful companies fall among the lowest ranked companies.

From the customer perspective, we see that the three successful companies (Amazon.com, ebay, and Priceline.com) are highly efficient (scores between .90 and 1.0), whereas two of the three failed companies (Furniture.com and PlanetRx.com) rank lowest on efficiency in the Customer perspective. The other failed company, Webvan falls in the middle range of efficiency. Apparently, these efficiency scores based on data from 1999 when all the companies were active, do differentiate between the ones that remained successful and those that subsequently failed. On the innovation and learning perspective, again the efficiency scores appear to discriminate between the subsequently successful and unsuccessful companies (Table 3). All three successful companies fall within the top 7 rank-ordered companies and have efficiency scores ranging from .79 to 1.0. On the other hand, the three companies that subsequently failed fall in the lowest six rank ordered companies, with efficiency scores less than.38.

On the internal process dimension, the results are mixed (as they were on the financial dimension). The two most successful companies (Amazon.com and ebay.com) were optimally efficient (1.0), while one failed company Furniture.com) also had high efficiency scores (.89). One successful and one failed company (Priceline.com and PlanetRx.com) had medium efficiency levels (.51-.68), and one failed company (Webvan) had the lowest efficiency score.

Table 3. Innovation & Learning Perspective and Internal Process Perspective

Innovation Perspective		Internal Process Perspective	
DMU Name	VRS Efficiency	DMU Name	VRS Efficiency
Amazon.com	1.00000	Amazon.com	1.00000
eBay	1.00000	Ebay	1.00000
Buy.com	1.00000	Buy.com	1.00000
CarsDirect.com	1.00000	Cdnow.com	1.00000
1-800-Flowers	1.00000	iown	1.00000
PetsMart.com	1.00000	iprint	1.00000
Priceline.com	0.79800	1-800-Flowers	1.00000
Outpost.com	0.77800	PetsMart.com	1.00000
iPrint	0.72561	Furniture.com	0.89217
Peapod	0.63112	Outpost.com	0.84993
Beyond.com	0.62029	Priceline.com	0.68200
Cdnow.com	0.38981	Beyond.com	0.58523
Furniture.com	0.38274	PlanetRX.com	0.51092
iOwn	0.25468	Peapod	0.49084
NextCard	0.21103	NextCard	0.44615
PlanetRX.com	0.19840	CarsDirect.com	0.18986
E*Trade	0.17813	E*Trade	0.17813
Webvan	0.15688	Webvan	0.08401

Comparison of Key Performance Indicators for successful and failed companies: In the next set of analyses, the three companies that remained successful were compared to the three companies that subsequently failed (filed for bankruptcy in 2000) on key performance indicators representing the four dimensions of the BSC (see Table 4). Although funding and revenues for the three successful companies are generally higher than for the three failed companies, five of the six companies show no profit (show negative profitability). Thus, the financial perspective does not present the complete picture. As can be seen from Table 4, two of the key performance indicators representing the Customer dimension (customer conversion factor and profitability per customer) appear to differentiate between the successful and failed companies, though results on revenue/customer is not quite as clear cut. All the key performance indicators for the Innovation dimension appear to differentiate between the successful and failed companies. This is similarly true for the Internal process dimension.

Table 4. Key Performance Indicators

	Finance Perspective			Customer Perspective		
DMU	Funding	Revenue	Profit	Conversion Factor	Rev./ Custmr	Profit./ Custmr
Amazon	2680.0	1640.0	-719.7	9.5	97	-42
eBay	823.9	224.7	10.8	5.9	22	1
Priceline	1592.0	482.4	-152.6	7.2	127	-40
Furniture	84.0	10.9	-46.5	2.4	42	-179
PlanetRX	144.5	8.9	-98.0	1.5	35	-386
Webvan	966.0	13.3	-144.6	2.8	283	-3077
	Innovation Perspective			Internal Process perspective		
	Rev./ Employee	Profit/ Employee	Cust./ Dev. Exp.	Conversion Factor	Rev./ Exp.	Rev./ Dev Exp.
Amazon	215789	-93680	106	9.5	4.0	10.3
eBay	741700	36093	421	5.9	2.3	9.4
Priceline	1276190	-403704	271	7.2	6.1	34.5
Furniture	51192	-218122	39	2.4	0.3	1.6
PlanetRX	23051	-251318	20	1.5	0.2	0.7
Webvan	47706	-518280	3	2.8	1.1	0.9

To summarize, all three non-financial key performance indicators do appear to differentiate between the successful and failed companies while the results of the financial perspective are mixed. Thus, the results show that the measures derived from the Balanced Scorecard framework can assist managers in predicting the success or failure of E-commerce companies.

References

Agrawal, V., Arjona L. D., and Lemmens R., "McKinsey B2C e-performance scorecard", *The McKinsey Quarterly*, 2001, Number 1.

Banker, R.D., Charnes, A. and Cooper, W.W., Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science.* **30**, 1078-1092 (1984).

Charnes, A., Cooper, W.W., and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research* 2/6, 429-444.

Kaplan, R. S. and D. P. Norton (1992), "The Balanced Scorecard – Measures that drive performance", *Harvard Business Review*, Jan-Feb., 71-79.

Seybold, P. B. and R. Marshak, Customers.Com: How to create a profitable business strategy for the internet and beyond, Random House Audio Books, 1999.

Zhu, Joe (2002), Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets. Kluwer Academic Publishers.

A Study of Variance of Locational Price in a Deregulated Generation Market

Jin-Tang Peng, Chen-Fu Chien

Department of Industrial Engineering and Engineering Management, National Tsing Hua University, 101 Sec. 2 Kung-Fu Rd. Hsinchu 300, Taiwan, R.O.C.

Abstract

This study investigates various scenarios with load variation and different bidding strategies by examining the locational prices in a deregulated power market that is modeled with a DC load flow approximation for minimizing production cost.
Key Words : Deregulation, Power Market, Congestion Management, Simulation

1. Introduction

Deregulation of power industry has become a worldwide trend, undergoing in many countries including Taiwan, to restructure the traditional monopoly power industry for introducing fair competition and improving economic efficiency. Power industry in Taiwan will be deregulated to take advantage of the operation efficiency on the basis of the competition of the market participants. However, the employment of competition in electric power markets in U.S. has not been completed [3]. As liberalization of the electric industry is becoming a reality in Taiwan, some questions regarding the ability of deregulated power market to provide an environment to openly trade electricity with operation efficiency and fair competition require urgent answers.

This study aims to explore the mechanism for power market by using scenario analysis and to thus propose the appropriate operation guides to avoid possible failure of electric power market. A pool-based market design through auction mechanism and considering locational prices is proposed. We construct a power delivery network as the example to examine different locational prices with load variation and bidding price variation in five different scenarios.

2. Proposed Market Mechanism

The auction system is an economically efficient mechanism to set prices and allocate demand to suppliers. However, its uniform auction price cannot provide locational price signals for the suppliers and consumers. Alternatively, nodal prices (i.e., locational marginal prices; LMP) can send signals to suppliers and consumers for allocating resources and demand. The effects of transmission constraints are reflected in the price at each node of the power network [1].

A pool-based market design that operates with an auction mechanism while considering nodal spot pricing prices of the network to reveal efficient signals to market participants was proposed. The pool model relies on dispatch actions of an Independent System Operator (ISO) to match the most efficient sources of power supply with customer demand while considering transmission constraints [2].

Multiple goals that have to be satisfied in the solution of proposed market model include the least production cost, the reliability of the system operation, and the balance of the regional demands and supplies. A DC power flow model is applied with the objective of the least-cost economic dispatch as follows:

$$Min \sum_{i \in I} P_i G_i \tag{1}$$

Subject to (a) generation constrains for all generators i:

$$G_i^{\min} \leq G_i \leq G_i^{\max} \quad or \quad G_i = 0 \quad , \quad \forall i \in I \tag{2}$$

(b) power flow equation for each bus j:

$$B\theta_j = G_{ij} - D_j \quad , \forall i \in I \ , \ \forall j \in J \qquad (3)$$

(c) line thermal limit constraint (i.e., transmission capacity) for all pairs j-k corresponding to existing lines:

$$\left| \frac{\theta_j - \theta_k}{X_{jk}} \right| \le P_{jk}^{\max} \quad , \forall j \in J \ , k \in J \ , j \ne k \qquad (4)$$

where P_i denotes the bid price of generator i, G_i denotes the quantity of power produced by the generator i, $G_i^{\min}$ denotes the minimum quantities that generator i can produce, and $G_i^{\max}$ denotes the maximum quantities that generator i can produce. In addition, B denotes the transfer admittance matrix (i.e., Linearised Jacobian matrix), θ_j denotes the phase angle of bus j, D_j denotes the load of bus j, X_{jk} denotes the impedance of line between bus j and k, and $P_{jk}^{\max}$ denotes the maximum power flow between bus j and k.

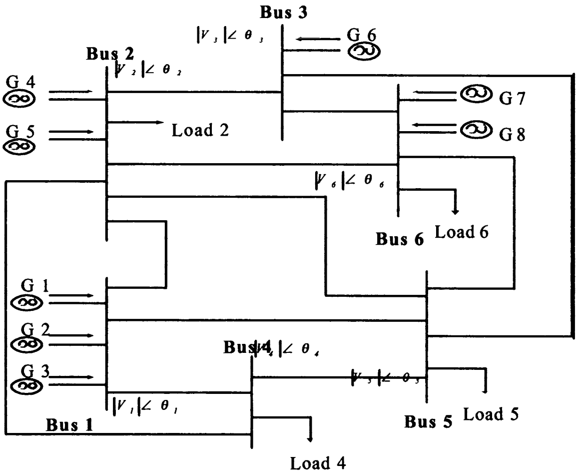


Fig. 1. Power systems of six buses

The transmission system differentiates the electric power market, where the flow of power cannot be easily controlled and the scarcity of transmission capacity leads to congestion [3]. The LMPs are calculated as dual variables (i.e., shadow price) for the equality constraint at the corresponding node to compute the optimal dispatch. That is, a LMP is the marginal cost of supplying the next increment of energy at a specific bus.

3. Scenario and Simulation Analysis

Considering the variance of load and generation cost, we simulate different bidding strategies of the generators and discuss the congestion cost of the transmission line and locational price of a power delivery network. A six-node network with eight generators (Gi, i=1,.., 8), six loads (Lj, j=1,.., 6), and eleven transmission lines was considered (Fig. 1). Table 1 summarizes the parameters including line thresholds, the capacities of generators, and the demands of loads.

Table 1. Parameters of generators, loads, and transmission lines

BUS	Plant	Minimum Volume[MW]	Maximum volume[MW]	Load [MW]
BUS1	G1	100	210	330
	G2	90	200	
	G3	110	380	
BUS2	G4	160	650	350
	G5	100	230	
BUS3	G6	120	410	180
BUS4				230
BUS5				190
BUS6	G7	90	200	250
	G8	110	390	

Line	P_{12}	P_{14}	P_{15}	P_{23}	P_{24}	P_{25}	P_{26}	P_{35}	P_{36}	P_{45}	P_{56}	$P_{24}+P_{36}$
Limit	100	100	100	80	150	120	110	90	110	110	80	260

With loss of generality, in scenario 1, we use a basic example in which the generator units offer the supply bids as shown in table 2. The ISO will determine the energy quantity and generators that are scheduled to run by minimizing the total cost subject to the constraints. In this case, Generators 1, 3, 4, 6, and 7 are scheduled to run as given in table 2.

In scenario 2, we simulate the market operation under price war among the generators in which the bidding price of the all generators is reduced to be half and each load decrease 50MW. The solution shows that Generators 1, 3, 4, and 6 are scheduled to run. The locational price at all buses in this condition is around 1110. The lower locational marginal price is owing to the lower price bid.

Table 2. Result of scenario 1

	G1	G2	G3	G4	G5	G6	G7	G8
Bidding price [$/MW]	2200	2230	2220	2150	2210	2180	2220	2240
Scheduled amounts	210	0	288.7	650	0	291.3	90	0
Locational Price	BUS1	BUS2	BUS3	BUS4	BUS5	BUS6		
	2220	2162.2	2180	2292.4	2201.2	2179		

In scenario 3, we simulate the market operation with increasing loads. It shows that Generators 1, 3, 4, 6, and 7 are scheduled to run. The locational price at buses 1, 2, 3, 4, 5, and 6 are 2220, 2150, 2180, 2662.8, 2264, and 2220, respectively. The locational price of bus 4 is significantly increased, because the transmission line capacity becomes unavailable. Thus, building the transmission line to meet unserved load is important when the loads are increase progressively.

In scenario 4, we simulate the market operation that the generator units all use high price strategies. It shows that all of the generators are scheduled to run except units 1 and 5. The locational price at buses 1, 2, 3, 4, 5, and 6 increase to be 2230, 2350, 2380, 2443.4, 2378.8, and 2240, respectively. Thus, ISO should prohibit the generators collectively control the price, especially when there are few players at the early stage of deregulation environment. Generator may also apply high price bid when the fuel price increased and thus increase the electricity cost.

In scenario 5, we simulate the condition of the transmission grid malfunction. The transmission grid is basic infrastructure and its stability is important. Suppose that transmission lines P_{24} and P_{25} diminish their capacities with 30MW, while the other conditions unchanged. It shows that the generators 1, 3, 4, 6, 7, and 8 are scheduled to run. The locational price at buses 1, 2, 3, 4, 5, and 6 are 2220, 2150, 2224.2, 2910.7, 2312.4, and 2220, respectively. The locational price of bus 4 is significantly increased because the transmission lines P_{14} and P_{24} reach their limits. The transmission system retains the feature of monopoly. Thus, the ISO who co-ordinates system operations should preserve its objectivity and reliability.

4. Discussion and Conclusion

On the supply side, the generator that adopts high price bid alone may not be scheduled to run, but if many of the generator units adopt the high price bids (e.g., collusion), the locational prices will be raised. The generator that adopts low price bid alone may increase its opportunity to be scheduled to run, however, as most of the generators use the low-price bidding strategies, the locational prices will be lowered and cause their profit loss even though they were scheduled to run. On the demand side, if the load grows, the probability of the congestion or potential overloading will be increased and thus increase the locational prices of the corresponding loads. Therefore, the generators may realize the revealed signal and further increase their bid prices. The constraints of the transmission network system operation affect the generator bidding and the power scheduling. The scarcity of transmission capacity leads to congestion and thus the locational prices and total system cost will be increased. Power industry with the traditional monopoly and have fewer participants in the supply and the transmission network. Therefore, the deregulation should start with increasing the generators and the available capacity of the transmission system. Small number of participants can collaboratively control the market. The transmission network should act as a common carrier that does not manipulate for private business interest. As the electrical industry becomes more competitive, transmission capacity constraints in production cost analysis and the accuracy of cost estimation is important.

References

1. Alvey T, Goodwin D, Ma X, Streiffert D, Sun D (1998) A security-constrained bid-clearing system for the New Zealand wholesale electricity market. IEEE Transactions on Power Systems, Vol. 13, No. 2: 340-346.
2. Finney JD, Othman HA, Rutz WL (1997) Evaluating transmission congestion constraints in system planning. IEEE Transactions on Power Systems, Vol. 12, No. 3: 1143-1148.
3. Silva C, Wollenberg BF, Zheng CZ (2001) Application of mechanism design to electric power markets. IEEE Transactions on Power Systems, Vol. 16, No. 1: 1-8.

Pseudo-Criterion Approaches to Evaluating Alternatives in Mangrove Forests Management

Santha Chenayah Ramu[1] and Eiji Takeda[2]

[1] Faculty of Economics and Administration, University Malaya, 50603 Kuala Lumpur, Malaysia
[2] Graduate School of Economics, Osaka University, Toyonaka, Osaka 560-0043, Japan

Abstract. The concept of the pseudo-criterion plays an important role in complex decision problems involving imprecise, uncertain and indeterminate data such as mangrove forests management. The outranking relation methods such as ELECTRE are well known to deal with the pseudo-criterion.

Ternary comparison method (TCM) in which a ternary AHP derives a priority vector from a single criterion concordance matrix is proposed to deal with a pseudo-criterion. Comparing the ranking of management alternatives for the mangrove forests by TCM with that by ELECTRE, we discuss the advantages and disadvantages of TCM.

1 Introduction

Sustainable use of mangrove forests requires a multiple criteria approach because mangroves are valuable natural resource with distinctive diversity, high intrinsic natural productivity and unique habitat value. So far, a few multicriteriaapproaches have been reported. Among them is Janssen and Padilla's work [1] which evaluated management alternatives of Pagbilao mangrove forests in the Philippines. To derive the ranking of alternatives, they used a simple weighted sum of standardized effects.

Since some values of alternatives are subjective indices and others are imprecise, and/or uncertain even if the criterion is quantitative, the pseudocriterion approach is preferable (see Roy and Vincke [2]). The outranking relation methods are used to deal with pseudo-criterion. Several versions of ELECTRE have been developed. Among others, ELECTRE III is the most familiar and has been widely used (see Roy [3] and Rogers et al. [4]). In general, it requires the weights assigned to criteria to build an outranking relation. It is, however, not uncommon that it is difficult to specify them precisely.

Ternary comparison method (TCM) [5] in which a ternary AHP [6] derives a priority vector from a single criterion concordance matrix incorporates a restricted set of unknown weights into the model. Demonstrating the use of TCM to mangrove forests management, we shall discuss the advantages and disadvantages of TCM.

2 Ternary Comparison Method (TCM)

Let us consider the set A of n alternatives:

$$A = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}.$$

Let $g_1, g_2, \dots, g_m$ be m-criteria. Thus, each alternative $\boldsymbol{a}_i$ is characterized by a multiattribute outcome denoted by a vector

$$(g_1(\boldsymbol{a}_i), g_2(\boldsymbol{a}_i), \dots, g_m(\boldsymbol{a}_i)).$$

In what follows, we assume that the decision maker prefers larger to smaller values for each criterion. Now, construct a ternary comparison matrix $T_k = (t_{ij}^k)$ for each criterion g_k:
(i) Strict preference $\boldsymbol{a}_i P \boldsymbol{a}_j$: $g_k(\boldsymbol{a}_i) - g_k(\boldsymbol{a}_j) > p_k$, then $t_{ij}^k = \theta$ and $t_{ji}^k = 1/\theta$
(ii) Weak preference $\boldsymbol{a}_i W \boldsymbol{a}_j$: $q_k < g_k(\boldsymbol{a}_i) - g_k(\boldsymbol{a}_j) \leq p_k$, then $t_{ij}^k = \theta$ and $t_{ji}^k = 1$
(iii) Indifference $\boldsymbol{a}_i I \boldsymbol{a}_j$: $| g_k(\boldsymbol{a}_i) - g_k(\boldsymbol{a}_j) | \leq q_k$, then $t_{ij}^k = t_{ji}^k = 1$
where p_k and q_k are respectively a preference and indifferent thresholds.

In TCM , a ternary AHP [6] is applied to derive a priority vector being a ratio scale from T_k. That is, the eigenvector $\boldsymbol{u}_k = (u_1^k, u_2^k, \dots, u_n^k)$ (whose elements are normalized so that the maximum value is equal to 1) associated with the maximum eigenvalue of T_k is adopted as a priority vector. And the aggregate value of each alternative is obtained by an additive weighting rule:

$$z_j = \sum_{k=1}^{m} w_k u_j^k,$$

where w_k is a weight for criterion k.

Let the indices of alternatives be renumbered in the descending order of $g_k(\boldsymbol{a}_i)$, that is,

$$g_k(\boldsymbol{a}_1) \geq g_k(\boldsymbol{a}_2) \geq \dots \geq g_k(\boldsymbol{a}_n).$$

Then, since the corresponding ternary comparison matrix $T_k = (t_{ij}^k)$ satisfies $t_{il}^k \geq t_{jl}^k$ for all l, it follows from a well known property of the maximum eigenvector that $u_i^k \geq u_j^k$.

3 Pseudo-Criterion Approaches to Mangrove Forests Management

Janssen and Padilla [1] evaluated the following eight management alternatives: $\boldsymbol{a}_1$ = Preservation (PR); $\boldsymbol{a}_2$ = Subsistence forestry (SF); $\boldsymbol{a}_3$ = Commercial forestry (CF); $\boldsymbol{a}_4$ = Aqua-silviculture (AS); $\boldsymbol{a}_5$ = Semi-intensive aquaculture (SA); $\boldsymbol{a}_6$ = Intensive aquaculture (IA); $\boldsymbol{a}_7$ = Commercial forestry and intensive aquaculture (CF/IA); and a_8 = Subsistence forestry and intensive aquaculture (SF/IA). Then, they considered three criteria, economic efficiency,equity, and environmental quality simultaneously. The effects of management alternatives are summarized in Table 1(Table 5 in [1]).

Table 1. Annual values of management alternatives for the Pagbilao mangrove forest [1].

	unit	PR	SF	CF	AS	SA	IA	CF/IA	SF/IA
Valued effects									
Subsistence forestry	1000pesos		349.73						189.34
Commercial forestry	1000pesos			415.84	217.77			229.00	
Fishponds	1000pesos				6724.2	22000.0	17000.0	6328.0	6328.0
Fish on site	1000pesos	163.05	158.63	158.63	122.58	8.14	8.14	40.00	40.00
Fish off site	1000pesos	1.94	1.88	1.88	1.46	0.09	0.09	0.28	0.28
Total value	1000pesos	164.99	510.24	576.35	7066.01	22008.23	17008.23	6597.28	6557.62
Other effects									
Emissions	tons/year				20.00	40.00	100.00	50.00	50.00
Soil accretion	cm/year	1.00	0.34	0.42	0.22	0.10	0.05	0.13	0.15
Biodiversity	index	1.00	0.72	0.52	0.16	0.22	0.09	0.24	0.44
Shore protection	index	1.00	0.37	0.15	0.15	0.13	0.07	0.15	0.15
Eco-tourism	index	0.80	1.00	0.38	0.18	0.14	0.08	0.21	0.30

The first five rows in table 1 are valued effects and are related to economic efficiency and equity criteria. Others are not valued and are related to environmental criterion. The performance of the alternatives on these three criteria is shown in Table 2(Table 7 in [1]). Environment is defined as an index combining effects on soil accretion, emissions, shore protection, biodiversity, and eco-tourism. The relative weight of biodiversity within this index is ten times the relative weight of each of other effects.

Table 2. Performance of the alternatives on three criteria [1].

	unit	PR	SF	CF	AS	SA	IA	CF/IA	SF/IA
Efficiency	1000pesos/year	165	510	576	7065	22000	17000	6588	6558
Equity	1000pesos/year	165	510	576	341.8	8	8	260	230
Environment	index	12.8	9.0	6.2	-17.9	-37.4	-98.9	-47.1	-45.0

To derive the ranking of alternatives, Janssen and Padilla used a weighted sum (WS) of standardized effects.

We will define threshold values: (1) economic efficiency: $p_1 = 6600, q_1 = 2200, v_1 = 22000$; (2) equity: $p_2 = 172.8, q_2 = 57.6, v_2 = 576$; (3) emissions: $p_3 = 30, q_3 = 10, v_3 = 100$, ; and $p_i = 0.3, q_i = 0.1, v_i = 1.0, (i = 4, 5, 6, 7)$, where (4) soil accretion, (5) biodiversity, (6) shore protection, and (7) eco tourism,

Then, priority vectors which are obtained by TCM are as follows:

$\boldsymbol{u}_1 = (0.417,\ 0.450,\ 0.450,\ 0.596,\ 1.000,\ 0.900,\ 0.596,\ 0.596),$
$\boldsymbol{u}_2 = (0.484,\ 0.900,\ 1.000,\ 0.760,\ 0.343,\ 0.343,\ 0.580,\ 0.580),$
$\boldsymbol{u}_3 = (1.000,\ 1.000,\ 1.000,\ 0.895,\ 0.542,\ 0.359,\ 0.490,\ 0.490),$
$\boldsymbol{u}_4 = (1.000,\ 0.775,\ 0.775,\ 0.600,\ 0.458,\ 0.458,\ 0.500,\ 0.500),$
$\boldsymbol{u}_5 = (1.000,\ 0.900,\ 0.692,\ 0.412,\ 0.453,\ 0.374,\ 0.492,\ 0.692),$
$\boldsymbol{u}_6 = (1.000,\ 0.844,\ 0.500,\ 0.500,\ 0.500,\ 0.451,\ 0.500,\ 0.500),$
$\boldsymbol{u}_7 = (0.899,\ 1.000,\ 0.652,\ 0.450,\ 0.450,\ 0.413,\ 0.496,\ 0.597).$

We conducted three cases according to Janssen and Padilla. In all cases, following them, weights are assumed to be: $w_5 = 10w_3, w_3 = w_4 = w_6 = w_7$.

Case 1. Efficiency, equity and environment are equally important.

WS: CF → SF → AS → SA → PR → CF/IA → SF/IA → IA

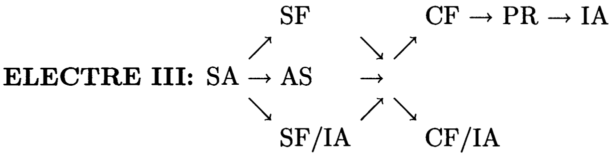
```
                      SA → AS → PR          IA
                    ↗                ↘    ↗
ELECTRE III:  SF                        CF
                    ↘                ↗    ↘
                         SF/IA               CF/IA
```

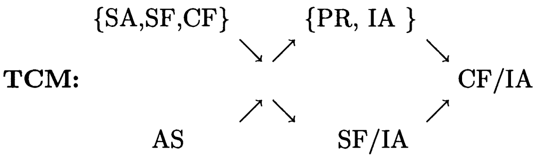
TCM: SF → CF → PR → AS → SF/IA → SA → CF/IA → IA
The restricted set of weights is: $W = \{\boldsymbol{w} \mid w_5 = 10w_3, w_3 = w_4 = w_6 = w_7$ $w_1 = w_2 = \sum_{i=3}^{7} w_i, w_i > 0, i = 1, 2, \ldots, 7\}$.

We find in all methods that IA and CF/IA are less preferred alternatives and that SF is ranked very highly. The results of CF are mixed. While CF is ranked very highly in WS and TCM, it emerges as the lower ranking in ELECTRE III.

Case 2. Efficiency is more important than equity, equity is more important than environment.

WS: SA → IA → AS → CF → SF → CF/IA → SF/IA → PR



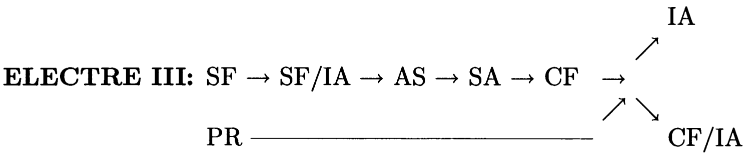


The set of weights is: $W = \{\boldsymbol{w} \mid w_5 = 10w_3, w_3 = w_4 = w_6 = w_7,$ $w_1 > w_2 > \sum_{i=3}^{7} w_i, w_i > 0, i = 1, 2, \ldots, 7\}$.

In case 2, SA, SF, and AS are ranked very highly in ELECTRE III and TCM. SF emerges as around the middle of the ranking in WS. While IA is the second best alternative in WS, it is the lower ranking in ELECTRE III and TCM.

Case 3: Environment is more important than equity, equity is more important than efficiency.

WS: CF → SF → PR → AS → SA → CF/IA → SF/IA → IA



TCM: SF → PR → CF → {AS, SF/IA} → {SA, CF/IA } → IA
The set of weights is: $W = \{\boldsymbol{w} \mid w_5 = 10w_3, w_3 = w_4 = w_6 = w7 > 0, w_1 < w_2 < \sum_{i=3}^{7} w_i, w_i > 0, i = 1, 2, \ldots, 7\}$.

In case 3, IA emerges as the least favored alternative and SF, PR are very favored alternatives in all methods. The results of CF are mixed. While CF is less favored in ELECTRE III, it is the best alternative in WS and the third best alternative in TCM.

4 Concluding Remarks

In the situations where imprecise and/or uncertain data exist together with qualitative criteria, a multiple pseudo-criterion approach is the best suited. TCM is a method in which a ternary AHP is adopted for a pseudo-criterion.

One of the difficulties in the outranking relation method such as ELECTRE III is that it requires the weights to specify the concordance index and uses a distillation method involving a certain amount of arbitrariness in the selection of a discrimination threshold function. One of the advantages of TCM is that it does not require the weights precisely. One of the distinguishing features of the outranking relation method is that it takes into account the discordance index by a veto threshold, while TCM does not. Since the procedures treating a pseudo-criterion necessarily involve arbitrariness, it may be preferable to derive the rankings of alternatives by different procedures for promoting complementary viewpoints.

References

1. Janssen, R., Padilla J. E. (1996): Valuation and Evaluation of Mangrove Alternatives for the Pagbilao Mangrove Forest, CREED Working Paper Series No.9, International Institute for Environment and Development, London, U.K.
2. Roy, B., Vincke, Ph. (1984): Relational Systems of Preference with One or More Pseudo-Criteria: Some New Concepts and Results. Management Science 30, 1323-1335
3. Roy, B. (1990): The Outranking Approach and the Foundations of ELECTRE Methods. In: Bana e Costa, C. A. (Ed.): Readings in Multiple Criteria Decision Aid. Springer-Verlag, Berlin Heidelberg,155-183
4. Rogers,M., Bruen, M., Maystre,L. Y. (2000): ELECTRE and Decision Support: Methods and Applications in Engineering and Infrastructure Investment. Kluwer Academic Publishers, Boston Dordrecht London
5. Takeda,E. (2001): A Method for Multiple Pseudo-Criteria Decision Problems. Computers & Operations Research 28, 1427-1439
6. Takahashi, I. (1990): AHP Applied to Binary and Ternary Comparisons. Journal of Operations Research Society of Japan 33,199-206

Energy-Environment-Cost Tradeoffs in Planning Energy Systems for an Urban Area

Hideharu Sugihara and Kiichiro Tsuji

Department of Electrical Engineering,
Graduate School of Engineering, Osaka University
2-1, Yamada-oka, Suita, Osaka 565-0871, Japan

Abstract. A multi-objective optimization model in which three objectives, i.e., primary energy consumption, CO_2 emission and cost are considered, has been developed for planning future energy systems in an urban area. The model is a mixed integer LP and it is useful to investigate a desirable energy system in small areas (2km by 2km) by generating a set of non-inferior solutions. The model has been applied to an actual city in Japan, and effectiveness of those relatively new, energy efficient or environmentally viable energy systems such as co-generation, fuel cell, and solar systems, will be described in terms of these tradeoffs.

1 Introduction

It is widely recognized as an urgent problem to reduce green house gases (CO_2, etc.) which cause the global warming issue. However, all countries do not perform sufficient countermeasures with diligence because people have to cease from current comfortable living and working space and economically efficient production systems to some extent. Also, in order to prevent from exhaustion of fossil energy resources, it is important to adopt new highly efficient energy technologies. Cost of the new technologies is still higher than those of conventional technologies. Thus, the issues of global warming, exhaustion of fossil energy resources and economic development can be essentially treated as a class of typical multi-objective optimization problems [1].

This paper proposes a new multi-objective optimization model for determining urban energy systems. The proposed model is applied to an actual city in Japan, and the tradeoff relation among three indices is analyzed.

2 Definitions of Energy System Alternatives

2.1 Customer Modeling

For simplicity, all customers in a specific area are grouped into seven types in the proposed model and they are geographically distributed in the area. These customer types are characterized by their daily end-use demand curves per floor space, such as space cooling and heating, and electricity (lighting, computer, etc.). Also, these seven customer types are divided into two sectors as follows;

- Business and Commercial sector : Office, Hotel, Hospital, Retail Store, Restaurant
- Residential sector : Detached House, Apartment

In the proposed model, a different set of energy system alternatives is designed for each sector as in the following section.

2.2 Energy System Alternatives (Individual type)

The energy system alternatives in business and commercial buildings, and residential houses are shown in Table 1 and Table 2, respectively. In Table 1, the energy systems "ARH" and "ER" mean gas absorption refrigerating and heater, and electric turbo refrigerator, respectively. These are widely used in current Japanese business and commercial buildings. Because "HP" system is equipped with heat accumulation for space heating and cooling demand, the "HP" system is effective for load leveling of electric power utility. Also, characteristics of FC system are lower in CO_2 emission and higher in equipment cost than GE system. Two operational policies for FC and GE are indicated by suffix 1 and 2. In Table 2, energy system "CNV" is considered to be widely used in recent Japanese residential houses. The "SLR1" and "SLR2" are solar utilization systems. The "SLR1" includes both solar power generation and solar water heater. The "SLR2" includes only the solar water heater and, therefore the "SLR1" has higher cost and lower CO_2 emission than "SLR2". Also, "ELE" depends only on electric energy supply. The "FC" is small-size fuel cell system for residential house, and its CO_2 emission is lower and its energy consumption is less than "CNV" system because the waste heat from the fuel cell is used for heated water supply. The "FC" is one of powerful options in future residential houses.

Table 1. Energy system alternatives in Business and Commercial sector

Symbols	Components
ARH	Absorption refrigerating and heating + Boiler
ER	Electric refrigerator + Boiler
HP	Heatpomp system + Heat accumuration equipment
GE1,GE2	Gas engine + Absorption refrigerator + Boiler
FC1,FC2	Fuel cell + Absorption refrigerator + Boiler
DHC	DHC(District Heating and Cooling)

Table 2. Energy system alternatives in Residential sector

Symbols	Components
CNV	Air-conditioner + Stove + Gas-boiler
SLR1	CNV + Solar power generation + Solar water heater
SLR2	CNV + Solar power generation
ELE	Air-conditioner + Electric water heater
DHC	DHC(District Heating and Cooling)

2.3 Centralized-type Alternative (District Heating and Cooling System)

The proposed model considers District Heating and Cooling (DHC) system as a centralized-type energy system. Cooling and heating energy (as cold and hot water) are supplied from a DHC plant to the customers included in the DHC covered area through piping networks as shown in Fig. 1. The DHC plant is a larger cogeneration plant that can be operated more efficiently than individual energy systems. As for piping network for transporting thermal energy, since the detailed network planning naturally becomes a large scale mixed integer optimization problem, and therefore, this paper assumes that each customer is directly connected to the DHC plant as shown in Fig. 1.

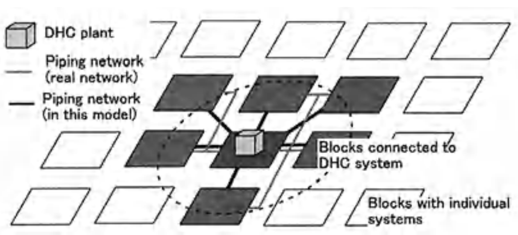


Fig. 1. Piping network model for DHC system

3 Formulation of multi-objective optimization model

3.1 Evaluation indices

The entire optimization problem is described as a mixed integer linear programming problem in which the share of each energy system alternative, the capacity and operational strategy of the DHC plant are determined simultaneously. The evaluating indices, i.e., cost, CO_2 emission and primary energy consumption, are calculated as follows:

$$
\begin{aligned}
J_{cost} &= \sum_{i \in I} CAP_i \cdot U_i + \sum_{n \in N} \nu_n \cdot \delta_n + D^E \cdot \overline{E} + D^G \cdot \overline{G} \\
&+ \sum_{s \in S} W_s \sum_{t \in T} \{ d^E_{s,t} \cdot x_{6,s,t} + d^G_{s,t} \cdot x_{7,s,t} + d^E sell_{s,t} \cdot x_{9,s,t} \\
&+ d^{GB}_{s,t} \cdot g^B_{s,t} + d^{GR}_{s,t} \cdot g^R_{s,t} \} + d^{oil} \cdot T_{oil} \\
&+ \sum_{m \in M_B} \sum_{j \in J_B} \{ \sum_{n \in N} A_{n,m} \} \cdot y^B_{m,j} \cdot CAP^B_{m,j} \\
&+ \sum_{m \in M_R} \sum_{j \in J_R} \{ \sum_{n \in N} A_{n,m} \} \cdot y^R_{m,j} \cdot CAP^R_{m,j} \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
J_{CO2} &= 0.0575 \cdot TX_7 + 0.1129 \cdot TX_6^{day} \\
&+ 0.0847 \cdot TX_6^{night} + 0.0805 \cdot T_{oil} \qquad (2)
\end{aligned}
$$

$$
J_{pri} = TX_7 + \frac{2450}{860} \cdot \{ TX_6^{day} + TX_6^{night} \} + T_{oil} \qquad (3)
$$

where(**variables**), U_i: capacity of equipment i in DHC plant, δ_n: block in DHC area (0-1 discrete variable), $\overline{E}, \overline{G}$: maximum demand of electricity and city-gas, $y^B_{m,j}, y^R_{m,j}$: non-negative share of energy system alternatives, $x_{6,s,t}$, $x_{7,s,t}$: purchased electricity and city-gas from outside utilities, $g^B_{s,t}, g^R_{s,t}$:city-gas consumption of individual energy systems at business and residential sector, TX_6^{day}, TX_6^{night}: annual purchasing elecric energy at day time and night time, TX_7, TX_{oil}: annual purchasing gas and kerosene, (**constants**), CAP_i: capacity cost of equipment i, ν_n: piping cost, D^E, D^G: charge for peak demand, $d^E_{s,t}, d^G_{s,t}$: electrcity and city-gas price for DHC plant, $d^{Esell}_{s,t}$: selling price of reverse power, $d^{GB}_{s,t}, d^{GR}_{s,t}$: city gas price for business and residential sector, d^{oil}: kerosene price, $A_{n,m}$: floor space, $CAP^B_{m,j}, CAP^R_{m,j}$: equipment cost per floor space of individual energy system, W_s: number of days, (**Suffix**), n: number of block, m: customer type, s: season, t: hour, i: equipment in DHC plant, j: energy system alternative, B: business and commercial sector, R: residential sector

In equation (2), the CO_2 emission coefficient of purchased electricity at day time and night time are different because of the different operating mix of electric power sources (coal, nuclear , oil and so on). Also, in equation (3), for conversion of electric energy to primary energy, coefficient is a reverse number of generating efficiency of outside electric utilities.

3.2 Constraints

Detail description of constraints is ommited due to the lack of space. As an important constraint, share variables of energy system alternatives must be summed up to 100%. Also, the purchased energy, i.e. electricity, city-gas and kerosene, from outside utilities are defined in order to evaluate the above three indices. A commercial solver package (GAMS/Cplex) is used for solving the above mixed integer linear programming. By using the solver package, an optimal solution can be obtained within a few minutes.

4 Tradeoff analyses

4.1 Studied area and analysis method

The proposed model is applied to an actual city (about 2km by 2km) in Japan. The ratio of the residential sector is over 60% in terms of floor space as shown in Table 3. A reference (current) energy system is defined as in Table 3 as a basis for making comparisons. Also, this paper adopts the constraint method in order to obtain non-inferior solutions.

Sector	Customer type	Floor area [%]	Reference system
Business and commercial sector	Office	23.8	ARH(24.4%), ER(75.6%)
	Hotel	1.7	ER
	Hospital	1.6	ER
	Retail store	8.3	ER
	Restaurant	1.6	ER
Residential sector	Detached house	36.9	CNV
	Apartment	26.1	CNV

Table 3. Floor space and their reference system

4.2 Results

In Fig. 2, the solutions are illustrated in two dimensional plane using CO_2 reduction rate as a parameter. In case of no CO_2 constraint, the maximum reduction of primary energy consumption is about 30 % at the cost increase of about 65 % compared with the current system. Taking CO_2 emission constraint into consideration, cost index become worse in some cases, because more environmentally viable and more expensive energy systems are adopted owing to CO_2 emission constraint. And also, in case of 20% reduction of CO_2 emission, all Pareto optimal solutions can reduce primary energy consumption simultaneously, i.e., positive reduction rate of primary energy consumption. It follows from this that primary energy consumption and CO_2 emission is not necessarily be independent each other. In case of no CO_2 constraint, the variations of energy systems in business and commercial sector, and residential sector are shown in Fig. 3 and Fig. 4, respectively. In Fig. 3, the energy system varies from the mix of DHC, GE and ARH to the dominance of the DHC system and then to the mix of DHC and individual FC systems. And in Fig. 4, the energy system varies from the dominance of CNV and ELE to FC and then to the mix of SLR1 and SLR2.

5 Conclusion

This paper proposed a multi-objective planning model for energy systems in an urban area, and illustrated tradeoff relations among cost, CO_2 emission and primary energy consumption. Using this model, the energy system alternatives are characterized from the viewpoint of primary energy consumption.

References

1. K.Tsuji et.al.(1999) Distributed Autonomous Energy Systems Planning for Urban Area. Proceedings of New Energy Systems and Conversions (NESC'99), Osaka, Japan, pp.455-460

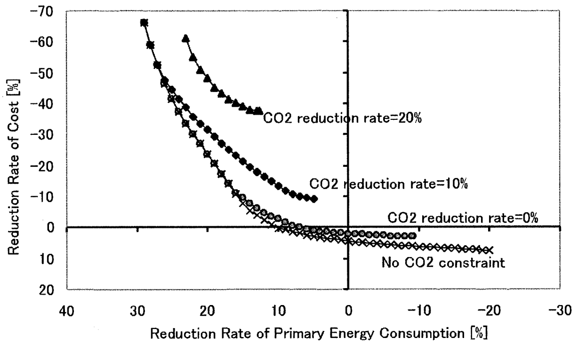


Fig. 2. Tradeoff curves

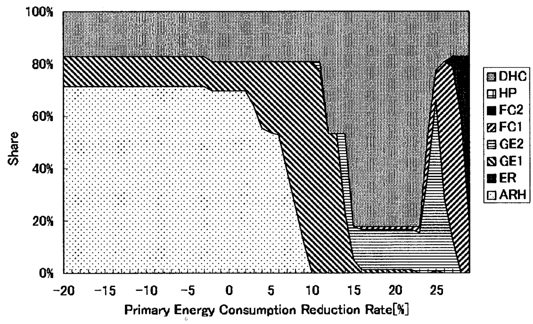


Fig. 3. Energy systems in business sector(no CO_2 constraint)

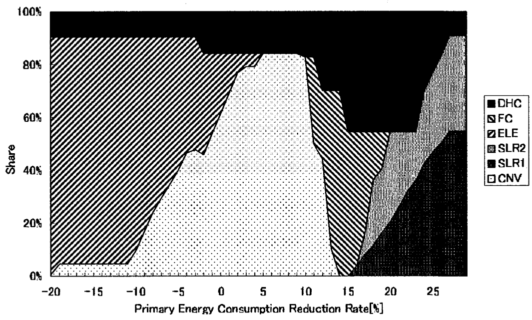


Fig. 4. Energy systems in residential sector(no CO_2 constraint)

DEA Approach to the Allocation of Various TV Commercials to Dayparts

Katsuaki Tanaka[1] and Eiji Takeda[2]

[1] Faculty of Business Administration and Information, Setsunan University, Neyagawa, Osaka 572-8508, Japan
[2] Graduate School of Economics, Osaka University, Toyonaka, Osaka 560-0043, Japan

Abstract. This paper presents an application of DEA to *ex ante* planning based on an empirical data, while many applications of DEA are directed to the *ex post* evaluations of accomplishments of competing DMUs.

For TV advertising planning, it is crucial to identify best practice commercials. However, though much effort has been done on determining advertising-response curves, there is little conclusive evidence as to the functional form of the maximum awareness value of the audiences given the repetitive insertions of a certain TV commercial to several dayparts.

We propose a simple non-parametric approach to an optimization problem with unknown functional form of the frontier based on DEA. Using the awareness data of TV commercials by TV-CM KARTE as past observations, an illustrative application to the allocation of various TV commercials to dayparts is provided.

1 Introduction

There is no doubt that TV commercial is very attractive to advertisers. But we have to concede that TV commercial is costly. Thus, how to make TV commercial more efficient is an important theme for advertisers.

To to this, it is crucial to identify the best practice commercials. For the allocation problem in TV advertising planning, however, it is difficult to specify the functional form of the maximum awareness value of the target audiences given the repetitive insertions of a certain TV commercial to several dayparts. In such cases, the parametric approach which has been most frequently used may not be appropriate.

Instead, an efficient frontier is estimated from sample data using DEA. Banker [1] showed that DEA estimator of the production frontier is the ML estimator under appropriate conditions and that the frontier estimator is biased below the theoretical frontier for a finite sample size and that asymptotically this bias reduces to zero.

Since its introduction, many applications of DEA have been reported (see Charnes et al. [2], Cooper et al. [3]). In marketing applications, for instance, Horsky and Nelson [4] discussed the relationship between the size of sales personnel and sales response and the reallocation of the sales personnel using a bench-marking technique based on DEA. Mahajan [5] presented a DEA model for assessing the relative efficiency of sales units that incorporate multiple and/or conflicting resources and outcomes.

This paper proposes a simple DEA model for allocating various TV commercials among the dayparts based on an empirical data. An illustrative application in this paper is based on the awareness data collected from a sample of 56 commercials in 1997 by TV-CM KARTE [6] which include both the original awareness values in all age levels and the insertion number per commercial in each daypart. A column added on the right of the table in the appendix is the BCC efficiency score of each commercial (see Appendix).

2 DEA approach to the allocation of various TV commercials to dayparts

In general, television-broadcasting time is divided into several types that are termed as dayparts during each broadcast day. We consider four main dayparts often used in real world for analysis. That is, D_1 =A time, D_2 =Special B (SB)time, D_3 =B time, and D_4 =C time which are ordered according to unit costs, that is, A time class has the highest unit cost.

A target segment is defined in terms of demographics. According to TV-CM KARTE, eight market segments are considered, that is, M_1 = male aged 13-19, M_2 = male aged 20-34, M_3 = male aged 35-49, M_4 = male aged 50-59, M_5 = female aged 13-19, M_6 = female aged 20-34, M_7 = female aged 35-49, M_8 = female aged 50-59.

Let $C_1, C_2, \ldots, C_{56}$ be different commercials as past observations by TV-CM KARTE. Each observed commercial C_i is characterized by an input vector $\boldsymbol{x}_i = (x_{1i}, x_{2i}, x_{3i}, x_{4i})$ and an output vector $\boldsymbol{y}_i = (y_{1i}, y_{2i}, \cdots, y_{8i})$, where x_{ji} is the number of insertions of Commercial C_i within daypart D_j and y_{ki} be the "CM"-awareness value (%) of the commercial C_i with market segment M_k. For simplicity, we assume that x_{ji} is continuous. The fractional solutions shown in the results should be appropriately rounded up to the integer value in practice since the number of insertions in any daypart is integral. "CM"-awareness value is defined as the percentage of the number of individuals who answered "I watched that TV commercial" within the sample in each market segment.

Since, in general, the advertiser who has various TV commercials with different target audiences purchases a sufficiently large number of units in dayparts at once, let S_j be the number of units in daypart D_j available for the advertiser.

Assume that the advertiser has two TV commercials of $C^{(1)}$: young female cosmetics and $C^{(2)}$: senior citizen cosmetics to be inserted in TV repeatedly during a specified time period.

Consider how to allocate $\boldsymbol{S} = (S_1, S_2, S_3, S_4)$ to $C^{(1)}$ and $C^{(1)}$ so as to maximize each awareness value as much as possible. To this end, the compromise programming is employed to get a solution that is as close to the ideal point as possible. To measure the distance between each solution and the ideal point, it employs l_p -norm. When $p = 1$ or ∞, the problem becomes a linear programming problem (see Yu [7]).

Let $W_1 = W_2 = 1$ and target population weights be respectively
$\boldsymbol{w}_1 = (w_{t1}) = (0,\ 0,\ 0,\ 0,\ 0.5,\ 0.4,\ 0.1,\ 0)$
and
$\boldsymbol{w}_2 = (w_{t2}) = (0,\ 0,\ 0,\ 0.5,\ 0,\ 0,\ 0,\ 0.5)$.
Given a set of observed data $\{(\boldsymbol{y}_j, \boldsymbol{x}_j), j = 1, 2, \ldots, 56\}$, the compromise programming problem for a data envelopment approximation problem is then represented by:

$$\min\{\sum_{k=1}^{2} |z_k^* - z_k|^p\}^{1/p}$$

s.t

$$z_k = \sum_{t=1}^{8} w_{tk} y_t^{(k)}, k = 1, 2,$$

$$\boldsymbol{y}^{(k)} = \sum_{j=1}^{56} \lambda_j^{(k)} \boldsymbol{y}_j, k = 1, 2,$$

$$\boldsymbol{x}^{(k)} = \sum_{j=1}^{56} \lambda_j^{(k)} \boldsymbol{x}_j, k = 1, 2,$$

$$x_i^{(1)} + x_i^{(2)} \leq S_i,\ i = 1, 2, 3, 4,$$

$$\sum_{j=1}^{56} \lambda_j^{(k)} = 1, k = 1, 2,$$

$$\lambda_j^{(k)} \geq 0,\ k = 1, 2; j = 1, 2, \ldots, 56$$

where z_1^*, z_2^*, are ideal points, $\boldsymbol{y}^{(k)} = (y_1^{(k)}, y_2^{(k)}, \ldots, y_8^{(k)})$ and
$\boldsymbol{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_4^{(k)}), k = 1, 2$

Note that the obtained *best practice* allocation may be a virtual campaign which is constructed by a convex combination of actual past campaigns of which the set is called a reference set. One of the advantages of DEA model is that it provides not only the optimum allocation (a virtual campaign) but also its reference set. The reference set consists of the actual campaigns that are most like the *best practice* virtual campaign. Therefore it serves to locate the best practice as benchmarking.

Assume that $\boldsymbol{S} = (S_1, S_2, S_3, S_4) = (40, 40, 40, 40)$
Then an ideal point is $\boldsymbol{z}^* = (z_1^*\ z_2^*) = (71.65\%,\ 49.67\%)$

For l_1, we have a compromise solution: $z_1 = 71.0\%$, $z_2 = 27.6\%$, $\boldsymbol{x}^{(1)} = (17.3, 29.1, 23.4, 20.0)$, $\boldsymbol{x}^{(2)} = (22.7, 10.9, 16.6, 20.0)$ which are attained at $\lambda_1^{(1)*} = 0.680$, $\lambda_{30}^{(1)*} = 0.320$, $\lambda_7^{(2)*} = 0.065$, $\lambda_{15}^{(2)*} = 0.066$, $\lambda_{46}^{(2)*} = 0.368$,

$\lambda_{54}^{(2)*} = 0.502$, and $\lambda_j^{(i)} = 0$, otherwise. Thus, to achieve these awareness values, it may be useful to examine C_1 and C_{30} for $C^{(1)}$ and C_7, C_{15}, C_{46} and C_{54} for $C^{(2)}$ as benchmarking.

For l_∞, we have $z_1 = 59.5\%$, $z_2 = 37.5\%$, $\boldsymbol{x}^{(1)} = (18.6, 20.4, 23.2, 19.3)$, $\boldsymbol{x}^{(2)} = (21.4, 19.6, 16.8, 20.7)$ which are attained at $\lambda_{30}^{(1)*} = 0.257$, $\lambda_{35}^{(1)*} = 0.687$, $\lambda_{41}^{(1)*} = 0.056$, $\lambda_7^{(2)*} = 0.023$, $\lambda_{15}^{(2)*} = 0.442$, $\lambda_{46}^{(2)*} = 0.238$, $\lambda_{54}^{(2)*} = 0.297$, and $\lambda_j^{(i)} = 0$, otherwise.

Comparing the result for l_1 with that for l_∞, we can see that SB daypart is effective for both $C^{(1)}$ and $C^{(2)}$ but is more effective for $C^{(1)}$. The awareness of sinior citizen cosmetics can raise at the expense of SB daypart toward young female cosmetics.

3 Concluding remarks

We have proposed a simple non-parametric approach to the allocation problem of various commercials to dayparts based on DEA. As an illustrative example, its application to the *ex ante* planning was provided using the awareness data of commercials by TV-CM KARTE.

Limitations of our simple model are:

(1) A day is divided in 4 time classes, A,SB,B,and C. Therefore, various spots, which may affect TV-awareness differently, are classified into the same daypart.

(2) Campaign periods vary among commercials though too-long campaigns periods are excluded from the analysis.

(3) Pattern of a daypart combination is not taken into account, e.g., some campaigns are scheduled intensively in a short time period, others are scheduled extensively in a campaign period.

As was seen in an illustration, one of the advantages of the data envelopment approximation model is that it locates the reference set as well as the optimum virtual campaign. It is useful to examine the actual patterns of TV campaigns included in the reference set as benchmarking for further analysis. Thus, despite its simplicity, we can gain insights about the patterns of the combination of dayparts.

References

1. Banker,R. D.(1993): Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. Management Science 39,1265-1273
2. Charnes,A., Cooper,W. W.,Lewin,A. Y., Seiford,L. M.(1994): Data Envelopment Analysis, Theory, Methodology and Applications. Kluwer Academic Publishers, Boston Dordrecht London
3. Cooper, W. W., Seiford,L. M.,Tone, K. (2000): Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software, Kluwer Academic Publishers, Boston Dordrecht London

4. Horsky,D., Nelson,P.(1996): Evaluation of salesforce size and productivity through efficient frontier benchmarking. Marketing Science 15, 301-320
5. Mahajan,J.(1991): A data envelopment analytic model for assessing the relative efficiency of the selling function. European Journal of Operational Research 53,189-205
6. Video Research Ltd.(1997): TV-CM KARTE, Vol. 98.
7. Yu,P. L.(1985): Multiple-Criteria Decision Making: Concepts, Techniques, and Extensions. Plenum Press, New York London

Appendix

Table 1 TV-CM CARTE

C_j	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	BCC score
1	16	24	26	29	60	48	35	28.1	68.4	73.1	54.2	36.5	1.00
2	15	72	31	58	72.5	54.5	44.7	36.8	50	70.4	59.4	47.6	1.00
3	92	114	59	75	52.5	43.1	37.9	31.6	63.2	53.7	29.2	17.5	0.27
4	45	88	101	81	80	52.8	45.6	35.1	78.9	74.1	63.5	44.4	0.54
5	20	26	73	30	62.5	41.5	40.8	29.8	42.1	61.1	58.3	38.1	1.00
6	30	62	12	39	60	60.2	41.7	26.3	55.3	66.7	46.9	23.8	0.91
7	16	8	88	2	10	8.1	13.6	8.8	15.8	25.9	25	30.2	1.00
8	32	32	59	10	50	43.9	27.2	14	47.4	55.6	45.8	27	0.83
9	96	69	39	64	37.5	26	27.2	15.8	50	54.6	31.3	19	0.29
10	26	28	52	26	55	24.4	20.4	5.3	39.5	35.2	18.8	12.7	0.71
11	17	33	8	34	55	50.4	30.1	24.6	65.8	53.7	44.8	19	1.00
12	111	150	62	31	80	63.4	56.3	35.1	89.5	93.5	84.4	76.2	1.00
13	27	39	41	25	47.5	34.1	19.4	12.3	65.8	65.7	51	36.5	0.69
14	50	78	21	13	82.5	65	55.3	26.3	89.5	88.9	66.7	34.9	1.00
15	20	32	20	22	70	57.7	65	47.4	57.9	78.7	63.5	57.1	1.00
16	34	57	53	46	40	29.3	19.4	10.5	36.8	50.9	27.1	14.3	0.39
17	31	80	45	62	75	65	37.9	26.3	81.6	74.1	47.9	28.6	0.74
18	31	56	23	39	12.5	5.7	2.9	0	7.9	8.3	6.3	6.3	0.48
19	37	36	11	14	70	56.9	41.7	29.8	60.5	81.5	45.8	27	1.00
20	79	114	30	39	75	61.8	66	50.9	71.1	84.3	79.2	74.6	1.00
21	24	30	19	49	27.5	7.3	7.8	3.5	23.7	24.1	9.4	4.8	0.54
22	114	138	70	89	20	30.9	26.2	31.6	13.2	25	9.4	15.9	0.20
23	28	80	40	71	40	37.4	27.2	24.6	39.5	38.9	29.2	23.8	0.48
24	15	36	27	25	22.5	22.8	28.2	19.3	18.4	50.9	47.9	38.1	0.99
25	93	165	92	134	72.5	65	46.6	26.3	89.5	84.3	53.1	42.9	0.44
26	26	43	10	94	50	42.3	21.4	7	39.5	32.4	18.8	11.1	0.68
27	27	40	63	32	32.5	19.5	22.3	7	44.7	56.5	45.8	20.6	0.58
28	48	42	18	23	75	65	55.3	50.9	68.4	79.6	59.4	52.4	1.00
29	26	37	70	61	70	52	48.5	19.3	63.2	56.5	55.2	31.7	0.84
30	20	40	18	1	70	58.5	44.7	12.3	84.2	71.3	50	33.3	1.00

Table 1 TV-CM CARTE (continued)

C_j	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	BCC score
31	42	47	98	69	52.5	51.2	53.4	50.9	47.4	59.3	55.2	61.9	0.94
32	22	22	36	54	65	44.7	45.6	35.1	52.6	65.7	54.2	31.7	1.00
33	70	77	17	44	80	71.5	60.2	31.6	65.8	77.8	61.5	42.9	0.89
34	66	121	148	91	32.5	26	38.8	15.8	39.5	58.3	50	38.1	0.24
35	17	13	26	27	52.5	33.3	33	21.1	50	60.2	45.8	22.2	1.00
36	34	65	38	64	17.5	20.3	11.7	7	39.5	56.5	33.3	19	0.45
37	39	53	48	84	22.5	16.3	4.9	0	13.2	11.1	4.2	3.2	0.30
38	46	128	64	53	17.5	11.4	19.4	17.5	23.7	33.3	25	27	0.25
39	20	19	18	18	2.5	10.6	14.6	7	2.6	10.2	13.5	9.5	0.73
40	39	51	15	34	87.5	78	63.1	50.9	63.2	77.8	70.8	54	1.00
41	31	22	12	8	47.5	47.2	29.1	15.8	60.5	57.5	40.6	23.8	1.00
42	67	95	31	28	90	78.9	80.6	73.7	71.1	86.1	71.9	61.9	1.00
43	52	105	64	95	90	73.2	51.5	28.1	89.5	85.2	68.8	27	1.00
44	29	28	53	54	52.5	41.5	20.4	14	50	52.8	36.5	25.4	0.65
45	37	63	19	107	40	39	35	14	39.5	57.4	41.7	23.8	0.51
46	9	17	15	5	17.5	8.9	10.7	17.5	13.2	18.5	17.7	23.8	1.00
47	20	57	82	61	85	74	67	52.6	81.6	82.4	87.5	71.4	1.00
48	23	53	26	16	47.5	29.3	32	12.3	31.6	46.3	35.4	19	0.65
49	32	61	46	129	90	74	58.3	24.6	89.5	88	67.7	44.4	1.00
50	71	87	39	37	77.5	61	64.1	49.1	81.6	75.9	84.4	61.9	1.00
51	17	27	48	38	25	19.5	34	22.8	23.7	36.1	45.8	42	0.92
52	30	27	52	16	70	43.1	34	31.6	81.6	66.7	52.1	42.9	1.00
53	18	11	4	46	37.5	26.8	16.5	7 31.	6	32.4	16.7	7.9	1.00
54	34	4	8	33	35	30.9	30.1	24.6	26.3	33.3	34.4	36.5	1.00
55	71	117	24	118	57.5	65.9	59.2	50.9	71.1	72.2	58.3	46	0.82
56	37	59	17	5	77.5	58.5	38.8	19.3	92.1	71.3	54.2	33.3	1.00

Analyzing Alternative Strategies of Semiconductor Final Testing

Hung-Ju Wang, Chen-Fu Chien, and Chung-Jen Kuo

Department of Industrial Engineering & Engineering Management, National Tsing Hua University, 101 Sec. 2 Kung Fu Rd., Hsinchu 300, Taiwan, R.O.C.

Abstract

This study integrates cost management and statistical decision analysis to construct a decision framework to determine an optimal alternative that balances multiple objectives of efficiency, cost, and risk. The framework was implemented in a testing firm and the experimental results showed its practical viability.
Key Words: Decision Analysis, Semiconductor Manufacturing, Final Test

1. Introduction

The various testing alternatives consisting of different setups and testing processes will affect the overkill and underkill rates of tested products. Overkill indicates that the product is truly good, but the test result is bad, i.e., false bad. Underkill indicates that the product is truly bad, but the test result is good, i.e., false good [1]. The overkilled products are rejected and thus result in the loss of the corresponding manufacturing cost. The underkilled products are sold and thus result in customer complaints and purchase returns. There is a trade-off between the overkill and underkill rates [2]. Little research has been done to analyze semiconductor final testing alternatives. This study aims to develop a framework for analyzing alternative testing strategies and processes, and thus determining the optimal testing alternative. An empirical study was conducted in a final testing factory, whose primary product is Mask Read-Only-Memory (Mask ROM).

2. Research Framework

We propose a research framework in which the throughput of testing process, the various risks, and the costs associated with overkilled and underkilled products are considered to determine the optimal decision as illustrated in Fig. 1. A testing alternative is defined by its receipt that consists of several testing processes and the corresponding thresholds. Performing the different testing alternative will affect the cost and productivity of the testing factory. Tested products are graded as pass or fail, by the comparing the tested parameters with the thresholds. Notably, the passed products consist of truly good and underkilled (i.e., false good) products. The failed products consist of truly defective and overkilled (i.e., false bad) products.

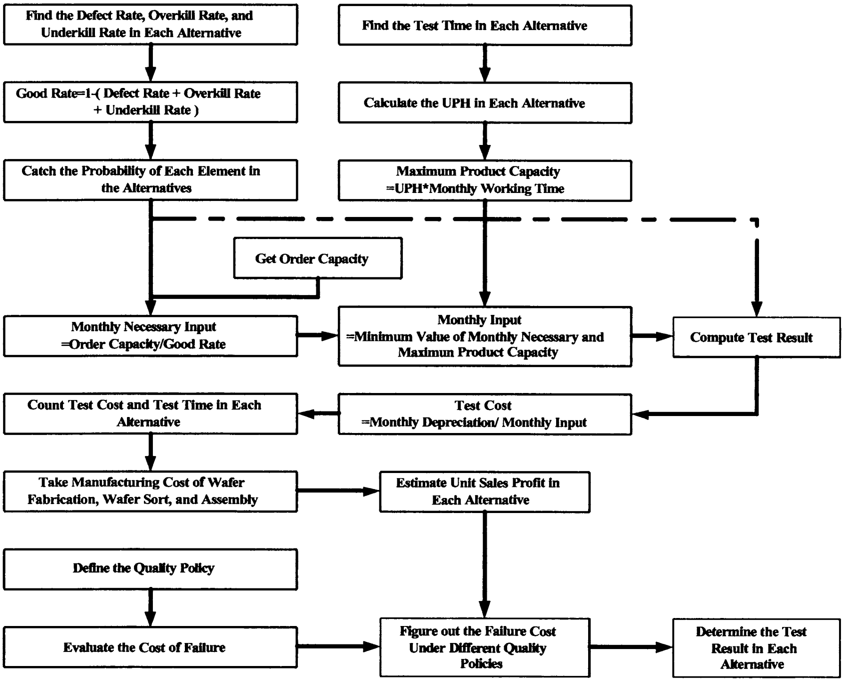


Fig. 1. Research framework

Let us first consider a test process that contains two sequential test procedures, A and B, as illustrated in Fig. 2, where o_A, u_A, and d_A denote the overkill rate, underkill rate and defect rate of test procedure A, respectively. Also, o_B, u_B, and d_B denote the overkill rate, underkill rate and defect rate of test procedure B, respectively. Following [3], in serial tests, the underkill rate becomes $u_B*(1\text{-}o_A\text{-}d_A)$.The total overkill rate will be $[o_A+o_B*(1\text{-}o_A\text{-}d_A)]$. Then, the defect rate of procedure B is derived by deducting its underkill rate from that of procedure A, i.e., $u_A - u_B*(1\text{-}o_A\text{-}d_A)$.Also, the true good rate of the two serial tests is derived by deducting the overkill rate of procedure B from the yield rate of procedure A, i.e., $(1\text{-} o_B)*(1\text{-}o_A\text{-}d_A)\text{-} u_A$.

Similar approaches may be easily extended to derive the parameters of a number of serial tests involved in different testing alternatives (see Fig. 3).

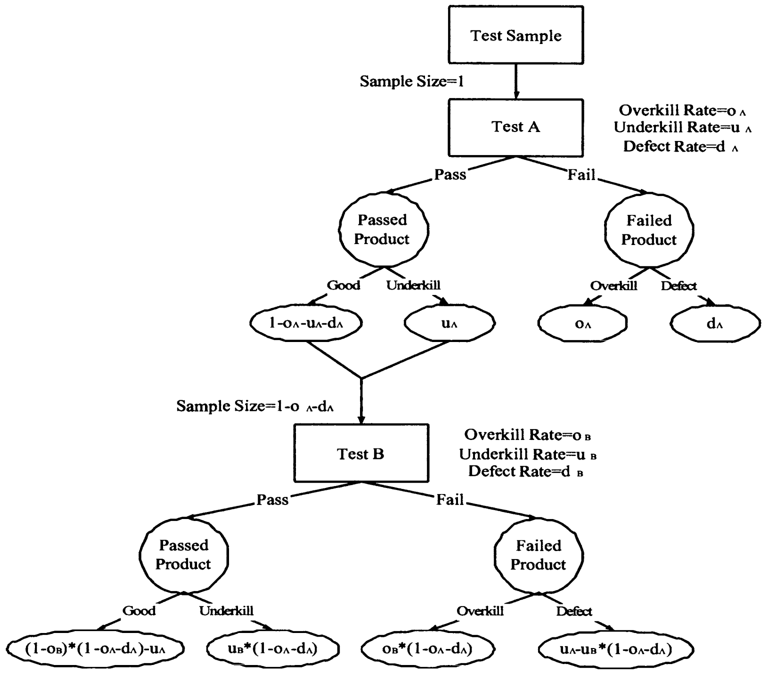


Fig. 2. Results of two sequential tests

As for the input data, the defect rate, overkill rate, underkill rate, and test time can be obtained from the historical or experimental data. The required monthly input can be calculated from both order capacity and yield rate. The throughput rate, *i.e.*, unit per hour (UPH), and the maximum product capacity can be calculated by testing time. Certainly, the monthly input is the lesser value between the monthly necessary input and the maximum product capacity. Then, the numbers of good, defect, overkill, and underkill products are calculated from the monthly input and the good rate, defect rate, overkill rate, and underkill rate. Based on monthly input, the test cost can be calculated with the monthly depreciation of the machine. The unit sale profit can be estimated with manufacturing cost of wafer fabrication, wafer sort and assembly. In addition, the quality policy disposes the product failure cost.

3. An Empirical Study

An empirical study was conducted in a semiconductor final testing firm in Taiwan. During hot seasons, the orders are nearly 30% larger than its capacities. To meet these orders and thus satisfy customer demand, the alternative of reducing testing time to increase capacity is investigated. The tested product has to pass both DC and AC test procedures. A high DC voltage input is used to ensure that electrostatic discharge (ESD) does not fail the product. An input operational AC impulse voltage is used to confirm that the stored data of this product is operable. There are mainly two approaches to reduce AC testing time: 1) reduces the AC test time in each test (Tcyc Reduction) and 2) reduces test procedures (Margin Mode).

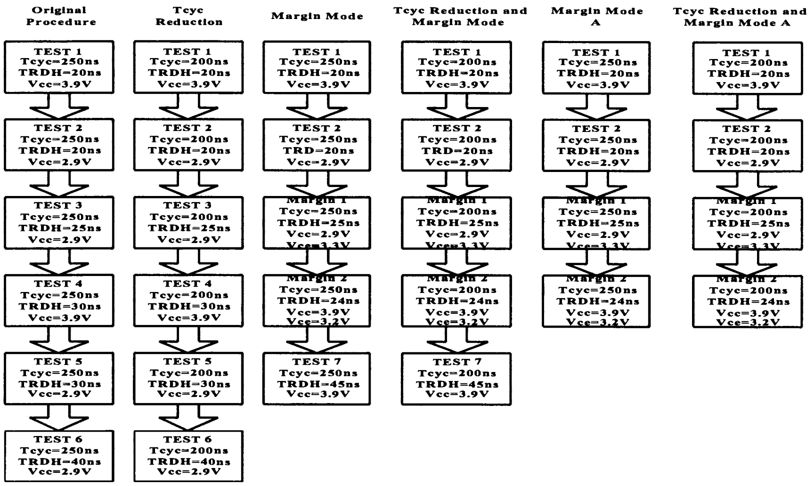


Fig. 3. Alternatives for Final Testing

In particular, we examined six alternative testing strategies as illustrated in Fig. 3. Tcyc denotes the cycle time in each AC test, TRDH represents the response time of MASK ROM during AC testing and Vcc or Vce are the input voltages of the AC test. The first alternative is the existing alternative that consists of six procedures. The second followed the existing procedures, but reduced the cycle time from 250 ns to 200 ns. The third and forth alternatives consisted of only five procedures, however, the latter reduces the cycle time from 250 ns to 200 ns. The fifth and sixth alternatives both consisted of four procedures while the sixth alternative also reduced the cycle time from 250 ns to 200 ns.

The defect rate, overkill rate, underkill rate and test time can be obtained from the historical and experimental data. From test data of the existing alternative, the corresponding underkill, overkill, and defect rates were derived. Although Tcyc reduction test reduced cycle time, it increased the overkilled rate. Meanwhile, for the test procedure with an input high AC voltage signal (i.e., AC-Hi, Vcc=3.9V), the overkill rate was also increased. Furthermore, because of increased overkill rate and relaxed product quality, the TRDH was increased from 20ns to 22ns. Indeed, experiments were conducted in two test batches that consist of 35 sets of underkill products. Each sample test batch contained 295 test products. Thus, the total overkill and underkill rates were derived as cycle time reduced. Table 1 summarizes the possible results of the six alternatives.

Table 1. Test results of the six alternatives

Test Alternatives	Original Procedure	Tcyc Reduction	Margin Mode	Margin Mode & Tcyc Reduction	Margin Mode A	Margin Mode A & Tcyc Reduction
Good Rate	94.656%	92.261%	94.461%	91.95%	94.454%	92.9354%
Underkill Rate	0.039%	0.049%	0.079%	0.070%	0.0856%	0.0846%
Defect rate	5.00%	5.00%	5.00%	5.00%	5.00%	5.00%
Overkill Rate	0.305%	2.690%	0.46%	2.98%	0.46%	1.98%

The cycle time of the original test alternative was 18.1 seconds. The total cycle times of the other alternatives are proportional to the involved test steps as shown in Table 2. We converted the cycle time into UPH based on the company's own formula. That is, UPH=(A*B)/((C+D)*(1+E)), where A is 3600 sec/hr, B denotes the number of simultaneously tested products, C denotes the test time, D denotes the interface time, and E denotes the interrupt rate. Table 2 summarizes the derived UPH, the associated test cost, and the corresponding capacity. In addition, because of the different yield rates, the necessary input to deliver the order of 520,000 passed ICs were derived as given in Table 2.

Table 2. Test Time and Cost of the Alternatives

Test Alternatives	Test Time (sec)	UPH (ea./hr)	Test Cost (NTD/ea.)	Capacity (ea./month)	Necessary Input (Order=520,000)
Original Procedure	18.10	604.15	3.83	449,488	549,357
Tcyc Reduction	14.48	718.78	3.22	534,772	563,618
Margin Mode	15.08	696.86	3.32	518,464	564,874
Margin Mode and Tcyc Reduction	12.07	882.69	3.06	612,081	565,525
Margin Mode A	12.07	882.69	3.13	612,081	550,530
Margin Mode A & Tcyc Reduction	9.65	962.41	3.07	716,033	559,528

Then, we derived the various profits and costs associated with the test alternatives. The sale price of each M-ROM was NT$132 dollars. The manufacturing cost of wafer fabrication, wafer sort, and assembly were $99.6, $1.9, and $6.9, respectively. An overkilled product that is considered to be defect will not be sold and thus cause a profit loss of its sale price. On the other hand, an underkilled product that is defect, yet has been sold will be identified by the customer later. According to the quality policy, which is the common, customers receive full refund of sale price, if they receive an underkilled (defect) product. Notably, the sale profit is the number of good products multiplied by the unit sale profit. The failure cost is the accumulation of defect cost, as well as overkill and underkill costs. The defect cost equaled the number of defects multiplied by the unit defect cost. The overkill cost equaled the number of overkills multiplied by the unit overkill cost. Similarly, the underkill cost equaled the number of underkills multiplied by the unit underkill cost. Table 3 presents the results of the above calculations, which are similar to the die level cost model. In particular, Margin mode A is the most profitable alternative in this case study.

Table 3. Profit Analysis of the Alternatives (Unit: NT$ 1,000)

Test Alternative	Original Procedure	Tcyc Reduction	Margin Mode	Margin Mode & Tcyc Reduction	Margin Mode A	Margin Mode A & Tcyc Reduction
Sales Profit	8,412	10,055	9,932	10,691	10,644	10,670
Defect Cost	2,522	2,985	2,896	3,151	3,070	3,118
Overkill Cost	181	1,899	315	2,225	334	1,462
Underkill Cost	23	35	54	52	62	62
Operation Income	5,685	5,137	6,667	5,263	7,178	6,027

4. Conclusion

In this study, a statistical decision analysis framework that tradeoffs multiple objectives of the different risks, the associated costs, and the efficiency (i.e., UPH) was developed to analyze various testing alternatives. An empirical study was conducted in an IC testing firm in Taiwan. The experimental results validated the practical viability of the proposed framework. Further research is required to collect related data and update the parameters continuously in the light of technology changes in IC products and testing equipment. A decision support system with the proposed framework embedded and the related data collected may be developed to assist such decisions to increase testing profits dynamically.

References

1. Montgomery DC (1991) Introduction to Statistical Quality Control, 2nd edition. John Wiley, New York.
2. Chien C, Hsu S, Peng C, Wu C (2000) A Cost-based Heuristic for Statistically Determining the Sampling Frequency in a Wafer Fab, IEEE SMTW Proceedings, 217-229.
3. Berger JO (1985) Statistical Decision Theory and Bayesian Analysis, 2nd edition. Springer-Verlag, New York.

A Discrete-Time European Options Model under Uncertainty in Financial Engineering

Yuji Yoshida

The University of Kitakyushu, 4-2-1 Kitagata, Kokuraminami, Kitakyushu 802-8577, Japan

Abstract. A discrete-time European options model with uncertainty of both randomness and fuzziness is presented, by introducing fuzzy logic to the stochastic financial model. The randomness and fuzziness in the systems are evaluated by both probabilistic expectation and fuzzy expectation, taking account of buyer's/writer's subjective demand goal. Fuzzy prices of European call/put options with uncertainty are given and their valuation and properties are discussed under a reasonable assumption. The meaning and properties of buyer's/writer's permissible range of expected prices are discussed in a numerical example.

1 Introduction

In mathematical modeling for European call/put options with uncertainty, the study of the discrete-time model is one of the most important approaches to investigate the continuous-time model through approximation (Pliska [3], Ross [5] and so on). We discuss the financial model with randomness and fuzziness as uncertainty from the viewpoint of fuzzy expectation, taking account of human subjective judgment. In this article, probability is applied as the uncertainty such that something occurs or not with probability, and fuzziness is applied as the uncertainty such that we cannot specify the exact values because of a lack of knowledge regarding the present stock market. By introducing fuzziness to stochastic processes in decision-making, we present a model with uncertainty of both randomness and fuzziness, which is a reasonable and natural extension of the original stochastic process. Next we discuss confidence intervals of European options prices with uncertainty (randomness and fuzziness) to secure the expected prices under investor's fuzzy goal which is considered as a demand function.

In order to describe a finance model with fuzziness, we need to extend real-valued random variables in the classical probability theory to *fuzzy random variables* which are random variables with fuzzy number values. Fuzzy random variables were first formulated mathematically by Puri and Ralescu [4] and have been studied by many authors. This article derives a recursive equation for the fuzzy option price process and gives a method to solve the problem without loss of worthy information contained in uncertainty like randomness and fuzziness.

2 Fuzzy stochastic processes

First, we give some mathematical notations regarding fuzzy random variables. Let $(\Omega, \mathcal{M}, P)$ be a probability space, where $\mathcal{M}$ is a σ-field of Ω and P is a non-atomic probability measure. $\mathbb{R}$ denotes the set of all real numbers. A fuzzy number is denoted by its membership function $\tilde{a} : \mathbb{R} \mapsto [0,1]$ which is normal, upper-semicontinuous, fuzzy convex and has a compact support. We identify fuzzy numbers with its corresponding membership functions. $\mathcal{R}$ denotes the set of all fuzzy numbers. The α-cut of a fuzzy number $\tilde{a}(\in \mathcal{R})$ is given by $\tilde{a}_\alpha := \{x \in \mathbb{R} \mid \tilde{a}(x) \geq \alpha\}$ for $\alpha \in [0,1]$ and $\tilde{a}_0 := \mathrm{cl}\{x \in \mathbb{R} \mid \tilde{a}(x) > 0\}$, where cl denotes the closure of an interval. We write the closed intervals as $\tilde{a}_\alpha := [\tilde{a}_\alpha^-, \tilde{a}_\alpha^+]$ for $\alpha \in [0,1]$.

An $\mathcal{R}$-valued map $\tilde{X}$ defined on Ω is called a fuzzy random variable if the maps $\omega \mapsto \tilde{X}_\alpha^-(\omega)$ and $\omega \mapsto \tilde{X}_\alpha^+(\omega)$ are measurable for all $\alpha \in [0,1]$, where $\tilde{X}_\alpha(\omega) = [\tilde{X}_\alpha^-(\omega), \tilde{X}_\alpha^+(\omega)] := \{x \in \mathbb{R} \mid \tilde{X}(\omega)(x) \geq \alpha\}$. Next we need to introduce expectations of fuzzy random variables. A fuzzy random variable $\tilde{X}$ is called integrably bounded if both $\omega \mapsto \tilde{X}_\alpha^-(\omega)$ and $\omega \mapsto \tilde{X}_\alpha^+(\omega)$ are integrable for all $\alpha \in [0,1]$. Let $\tilde{X}$ be an integrably bounded fuzzy random variable. The expectation $E(\tilde{X})$ of the fuzzy random variable $\tilde{X}$ is defined by a fuzzy number

$$E(\tilde{X})(x) := \sup_{\alpha \in [0,1]} \min\{\alpha, 1_{E(\tilde{X})_\alpha}(x)\}, \quad x \in \mathbb{R}, \tag{1}$$

where $E(\tilde{X})_\alpha := [E(\tilde{X}_\alpha^-), E(\tilde{X}_\alpha^+)]$ for $\alpha \in [0,1]$.

We consider a discrete-time *fuzzy stochastic process* defined by fuzzy random variables. Let T be a positive integer and let $\{\tilde{X}_t\}_{t=0}^T$ be a sequence of integrably bounded fuzzy random variables. $\{\mathcal{M}_t\}_{t=0}^T$ is a family of non-decreasing sub-σ-fields of $\mathcal{M}$ such that for $t = 0, 1, \cdots, T$ fuzzy random variables $\tilde{X}_t$ are $\mathcal{M}_t$-adapted, i.e. random variables $\tilde{X}_{r,\alpha}^-$ and $\tilde{X}_{r,\alpha}^+$ are $\mathcal{M}_t$-measurable for all $r = 0, 1, 2, \cdots, t$ and $\alpha \in [0,1]$. Then we call $(\tilde{X}_t, \mathcal{M}_t)_{t=0}^T$ a fuzzy stochastic process.

3 European options in uncertain environment

Let a positive real number r be an interest rate of a *bond price*, which is riskless asset, and let a discount rate $\beta = 1/(1+r)$. Let a positive integer T be an *expiration date*. Define a *stock price process* $\{S_t\}_{t=0}^T$ as follows: S_0 is a positive constant and

$$S_t := S_0 \prod_{s=1}^{t} (1 + Y_s) \quad \text{for } t = 1, 2, \cdots, T, \tag{2}$$

where $\{Y_t\}_{t=1}^T$ is a uniform integrable sequence of independent, identically distributed real random variables on $[r-1, \infty)$ such that $E(Y_t) = r$ for all $t =$

$1, 2, \cdots, T$. This condition means that there exists a risk-neutral measure, and then there is no arbitrage opportunity ([5]). The σ-fields $\{\mathcal{M}_t\}_{t=0}^T$ are given as follows: $\mathcal{M}_0$ is the completion of $\{\emptyset, \Omega\}$ and $\mathcal{M}_t$ denote the completions of σ-field generated by $\{Y_1, Y_2 \cdots Y_t\}$ for $t = 1, 2, \cdots, T$.

We consider a finance model where the stock price process $\{S_t\}_{t=0}^T$ takes fuzzy values. We give fuzzy values by triangular fuzzy numbers for simplicity. Let $\{a_t\}_{t=0}^T$ be an $\mathcal{M}_t$-adapted stochastic process such that $0 < a_t(\omega) \leq S_t(\omega)$ for almost all $\omega \in \Omega$. Then, *stock price process with fuzzy values* are represented by a fuzzy stochastic process $\{\tilde{S}_t\}_{t=0}^T$:

$$\tilde{S}_t(\omega)(x) := L((x - S_t(\omega))/a_t(\omega)) \tag{3}$$

for $t = 0, 1, 2, \cdots, T$, $\omega \in \Omega$ and $x \in \mathbb{R}$, where $L(x) := \max\{1 - |x|, 0\}$ for $x \in \mathbb{R}$ is the triangle-shape function and $\{a_t\}_{t=0}^T$ is a sequence of random variables with positive values. Hence, $a_t(\omega)$ is a spread of triangular fuzzy numbers $\tilde{S}_t(\omega)$ and corresponds to the amount of fuzziness in the process. The fuzziness in the process increases as $a_t(\omega)$ becomes bigger, and $a_t(\omega)$ should be an increasing function of the stock price $S_t(\omega)$ (see Assumption S in this section). The α-cuts of (3) are

$$[\tilde{S}_{t,\alpha}^-(\omega), \tilde{S}_{t,\alpha}^+(\omega)] = [S_t(\omega) - (1-\alpha)a_t(\omega), S_t(\omega) + (1-\alpha)a_t(\omega)]. \tag{4}$$

We define fuzzy stochastic processes of European call/put options by $\{\tilde{C}_t\}_{t=0}^T$ and $\{\tilde{P}_t\}_{t=0}^T$:

$$\tilde{C}_t(\omega) := e^{-rt}(\tilde{S}_t(\omega) - 1_{\{K\}}) \vee 1_{\{0\}} \tag{5}$$

$$\tilde{P}_t(\omega) := e^{-rt}(1_{\{K\}} - \tilde{S}_t(\omega)) \vee 1_{\{0\}} \tag{6}$$

for $t = 0, 1, 2, \cdots, T$, $\omega \in \Omega$, where $1_{\{K\}}$ and $1_{\{0\}}$ denote the crisp number K and zero respectively and $\vee$ is the maximum induced from the fuzzy max order ([2]): For fuzzy numbers $\tilde{a}, \tilde{b} \in \mathcal{R}$, the maximum $\tilde{a} \vee \tilde{b}$ is the fuzzy number whose α-cuts are $(\tilde{a} \vee \tilde{b})_\alpha = [\max\{\tilde{a}_\alpha^-, \tilde{b}_\alpha^-\}, \max\{\tilde{a}_\alpha^+, \tilde{b}_\alpha^+\}]$ for $\alpha \in [0, 1]$. The α-cuts of (5) and (6) are

$$\tilde{C}_{t,\alpha}(\omega) = [\max\{e^{-rt}(\tilde{S}_{t,\alpha}^-(\omega) - K), 0\}, \max\{e^{-rt}(\tilde{S}_{t,\alpha}^+(\omega) - K), 0\}]; \tag{7}$$

$$\tilde{P}_{t,\alpha}(\omega) = [\max\{e^{-rt}(K - \tilde{S}_{t,\alpha}^+(\omega)), 0\}, \max\{e^{-rt}(K - \tilde{S}_{t,\alpha}^-(\omega)), 0\}]. \tag{8}$$

We evaluate these fuzzy stochastic processes by the expectations introduced in the previous section. Then, the expectations of fuzzy price processes in European call/put options are given as follows:

$$\tilde{V}^C(y, t) := E(e^{-r(T-t)}(\tilde{S}_T - 1_{\{K\}}) \vee 1_{\{0\}} \mid S_t = y) \tag{9}$$

$$\tilde{V}^P(y, t) := E(e^{-r(T-t)}(1_{\{K\}} - \tilde{S}_T) \vee 1_{\{0\}} \mid S_t = y) \tag{10}$$

for an initial stock price y for $y > 0$ and $t = 0, 1, 2, \cdots, T$, where $E(\cdot)$ is the expectation with respect to some risk-neutral equivalent martingale measure

([3]). Put their α-cuts by $[\tilde{V}_\alpha^{C,-}(y,t), \tilde{V}_\alpha^{C,+}(y,t)]$ and $[\tilde{V}_\alpha^{P,-}(y,t), \tilde{V}_\alpha^{P,+}(y,t)]$ respectively.

We introduce a valuation method of fuzzy prices, taking into account of investor's subjective judgment. Let a fuzzy goal by a fuzzy set $\varphi : [0,\infty) \mapsto [0,1]$ which is a continuous and increasing function with $\varphi(0) = 0$ and $\lim_{x\to\infty} \varphi(x) = 1$. Then we note that the α-cut is $\varphi_\alpha = [\varphi_\alpha^-, \infty)$ for $\alpha \in (0,1)$. For an exercise time T and the call/put options with fuzzy values given in (5) and (6), we define a *fuzzy expectation* of the fuzzy numbers $\tilde{V} = E(\tilde{C}_T)$ or $\tilde{V} = E(\tilde{P}_T)$ by

$$\tilde{E}(\tilde{V}) := \fint_{[0,\infty)} \tilde{V}(x) \; \mathrm{d}\tilde{m}(x) = \sup_{x\geq 0} \min\{\tilde{V}(x), \varphi(x)\}, \tag{11}$$

where $\tilde{m}$ is the possibility measure generated by the density φ and $\fint \mathrm{d}\tilde{m}$ denotes Sugeno integral ([6]). The fuzzy number $\tilde{V} = E(\tilde{C}_T)$ or $\tilde{V} = E(\tilde{P}_T)$ means a fuzzy price, and the fuzzy goal $\varphi(x)$ represents the achievement degree of the buyer's/writer's demand prices x ([1]). Then, the fuzzy expectation (11) shows a degree of expected prices which is adequate for the investor's demand profits. Hence, a positive number x^* is called an expected price if it attains the supremum of the fuzzy expectation (11), i.e.

$$\tilde{E}(\tilde{V}) = \sup_{x\geq 0} \min\{\tilde{V}(x), \varphi(x)\} = \min\{\tilde{V}(x^*), \varphi(x^*)\}. \tag{12}$$

Now we introduce a reasonable assumption. We can develop the theory in this article without the following Assumption S and triangle-type shape functions (3). However this article adopts them for the numerical computation which is important for its application.

Assumption S. The stochastic process $\{a_t\}_{t=0}^T$ is represented by

$$a_t(\omega) := cS_t(\omega), \quad t = 0,1,2,\cdots,T, \; \omega \in \Omega, \tag{13}$$

where c is a constant satisfying $0 < c < 1$.

Assumption S is reasonable since $a_t(\omega)$ means a size of fuzziness and it should depend on the volatility and the stock price $S_t(\omega)$ because one of the most difficulties is estimation of the volatility in actual cases ([5, Sect.7.5.1]). In this model, we represent by c the fuzziness of the volatility, and we call c a *fuzzy factor* of the process. Further, since in an uncertain environment the final decision making should be done under investor's own subjective judgments, we adopt the fuzzy expectation which is the decision-maker's subjective estimation for the prices of options. From now on, we suppose that Assumption S holds. Then we have

$$\tilde{S}_{t,\alpha}^{\pm}(\omega) = S_t(\omega) \pm (1-\alpha)a_t(\omega) = b^{\pm}(\alpha)S_0 \prod_{i=1}^{t}(1+Y_i(\omega)), \quad \omega \in \Omega \tag{14}$$

for $t = 0, 1, \cdots, T$ and $\alpha \in [0,1]$, where $b^{\pm}(\alpha) := 1 \pm (1-\alpha)c$ for $\alpha \in [0,1]$. The following recursive results regarding the fuzzy prices in European call/put options are obtained by dynamic programming in a similar way to [1].

Theorem 1. *(Recursive equation).*

(i) *In European call option, it holds that*

$$\tilde{V}_{\alpha}^{C,\pm}(y,t) = \beta E(\tilde{V}_{\alpha}^{C,\pm}(y(1+Y_1), t+1)) \tag{15}$$

for $t = 0, \cdots, T-1$ *and* $y > 0$, *where* $\tilde{V}_{\alpha}^{C,\pm}(y,T) := \max\{b^{\pm}(\alpha)y - K, 0\}$.

(ii) *In European put option, it holds that*

$$\tilde{V}_{\alpha}^{P,\pm}(y,t) = \beta E(\tilde{V}_{\alpha}^{P,\pm}(y(1+Y_1), t+1)) \tag{16}$$

for $t = 0, \cdots, T-1$ *and* $y > 0$, *where* $\tilde{V}_{\alpha}^{P,\pm}(y,T) := \max\{K - b^{\mp}(\alpha)y, 0\}$.

4 The expected price of European options

In this section, we discuss the permissible ranges of the expected prices in European call/put options $\tilde{V} = \tilde{V}^C(y,0)$ or $\tilde{V} = \tilde{V}^P(y,0)$. Fix an initial stock price y. Define a grade $\alpha^{C,+} := \sup\{\alpha \in [0,1] \mid \varphi_{\alpha}^{-} \le \tilde{V}_{\alpha}^{C,+}(y,0)\}$, where $\varphi_{\alpha} = [\varphi_{\alpha}^{-}, \infty)(\alpha \in (0,1))$ and the supremum of the empty set is understood to be 0. We obtain the following theorem by modifying the results in Yoshida [7].

Theorem 2.

(i) *It holds that* $\alpha^{C,+} = \tilde{E}(\tilde{V}^C(y,0))$.
(ii) *The grade* $\alpha^{C,+}$ *satisfies* $\varphi_{\alpha^{C,+}}^{-} = \tilde{V}_{\alpha^{C,+}}^{C,+}(y,0)$.
(iii) *The corresponding expected price is given by* $x^{C,+} = \varphi_{\alpha^{C,+}}^{-}$.

Since the fuzzy expectation (11) is defined by possibility measures, $x^{C,+}$ gives an upper bound on expected prices of European call option. Therefore, similarly to Theorem 2 we can define another grade, which gives a lower bound on expected prices of European call option as follows:

$$x^{C,-} = \varphi_{\alpha^{C,-}}^{-}, \tag{17}$$

where $\alpha^{C,-}$ is defined by $\alpha^{C,-} := \sup\{\alpha \in [0,1] \mid \varphi_{\alpha}^{-} \le \tilde{V}_{\alpha}^{C,-}(y,0)\}$, and it satisfies

$$\varphi_{\alpha^{C,-}}^{-} = \tilde{V}_{\alpha^{C,-}}^{C,-}(y,0). \tag{18}$$

Hence, from (17), (18) and Theorem 2, we can easily check the interval

$$[x^{C,-}, x^{C,+}] = \{x \in \mathbb{R} \mid \tilde{V}^C(y,0)(x) \ge \varphi(x)\}, \tag{19}$$

which is the range of expected prices such that the degree of the expected price $\tilde{V}(x)$ is adequate for buyer's demand profits $\varphi(x)$. Therefore, the interval

$[x^{C,-}, x^{C,+}]$ means *the permissible range of the buyer's expected prices* for his demand φ. Regarding European put option, similarly we obtain writer's permissible range of expected prices by $[x^{P,-}, x^{P,+}]$, where $x^{P,-} := \varphi^-_{\alpha^{P,-}}$ and $x^{P,+} := \varphi^+_{\alpha^{P,+}}$ and the grades $\alpha^{P,-}$ and $\alpha^{P,+}$ are given by $\varphi^-_{\alpha^{P,-}} = \tilde{V}^{P,-}_{\alpha^{P,-}}(y,0)$ and $\varphi^-_{\alpha^{P,+}} = \tilde{V}^{P,+}_{\alpha^{P,+}}(y,0)$.

Example 1. We consider a binomial CRR-model (Ross [5, Sect.7.4]) to image the discrete-time European put option model presented in this article. Consider a fuzzy goal

$$\varphi(x) = \begin{cases} 1 - e^{-0.1x}, & x \geq 0 \\ 0, & x < 0. \end{cases} \tag{20}$$

Then $\varphi^-_\alpha = -0.1^{-1}\log(1-\alpha)$ for $\alpha \in (0,1)$. Put an exercise time $T = 8$, an interest rate of a bond $r = 0.05$, a fuzzy factor $c = 0.07$, an initial stock price $y = 20$, a strike price $K = 25$ and a volatility $\sigma = 0.25$. Let $p := (1 + r - e^{-\sigma})/(e^\sigma - e^{-\sigma})$. A sequence of random variables $\{Y_t\}_{t=1}^T$ is given by

$$Y_i := \begin{cases} e^\sigma - 1 & \text{with probability } p \\ e^{-\sigma} - 1 & \text{with probability } 1 - p \end{cases} \quad \text{for } t = 1, 2, \cdots, T. \tag{21}$$

Define a stock price process $\{S_t\}_{t=0}^T$ by (2) with (21). Then we can calculate that the grades of the fuzzy expectation of the fuzzy price are $\alpha^{P,-} \approx 0.89218$ and $\alpha^{P,+} \approx 0.89262$. The permissible range of the writer's expected prices in European put option under his demand φ is $[x^{P,-}, x^{P,+}] \approx [22.2729, 22.3136]$. Now we also calculate non-fuzzy option price $x \approx 22.2933$. Compare this with that fuzzy case. This difference is from the fuzziness of the process. Therefore, $[x^{P,-}, x^{P,+}]$ is a kind of confidence intervals of the option price under the investor's subjective judgment.

References

1. Bellman, R. E., Zadeh, L. A. (1970) Decision-making in a fuzzy environment. Management Sci. Ser. B. **17**, 141-164
2. Klir, G. J., Yuan, B., (1995) Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, London
3. Pliska, S.R. (1997) Introduction to Mathematical Finance: Discrete-Time Models. Blackwell Publ., New York.
4. Puri, M. L., Ralescu, D. A. (1986) Fuzzy random variables. J. Math. Anal. Appl. **114**, 409-422
5. Ross, S.M. (1999) An Introduction to Mathematical Finance. Cambridge Univ. Press, Cambridge.
6. Sugeno, M. (1974) Theory of fuzzy integrals and its applications. Doctoral Thesis, Tokyo Institute of Technology
7. Yoshida, Y. (1996) An optimal stopping problem in dynamic fuzzy systems with fuzzy rewards. Computers Math. Appl. **32**, 17-28.
8. Yoshida, Y. The valuation of European options in uncertain environment. Europ. J. Oper. Res., to appear.

Multipurpose Decision-Making in House Plan by Using AHP

Bingjiang Zhang[1], Hui Liang[2], and Tamaki Tanaka[1]

[1] Graduate School of Science and Technology, Niigata University,
Niigata 950-2181, Japan
E-mail: tamaki@m.sc.niigata-u.ac.jp
[2] Hirosaki University, Hirosaki 036-8561, Japan

Abstract. In this paper, we suggest and analyze a new support system for easy decision-making on housing planing. Using this, users can analyze their preferences by themselves individually and can select optimal alternatives. The purpose of this paper is to propose a rational decision-making system in housing planing based on the AHP (Analytic Hierarchy Process). Then, we show how standard evaluation can be decided on the this system. In this paper, evaluation and election of the room arrangements is based on the information data came from users.

1 Introduction

The basic problem of decision-making is how to choose best one in a set of competing alternatives that are evaluated under conflicting criteria. The method of Analytic Hierarchy Process (AHP) provides us with a comprehensive framework for solving such problems. Liang [1] has been reported by using the AHP to evaluate housing planing by qualitative elements such as several kinds of construction methods, room arrangement, and designs. Based on this research on room arrangements particularly, we observe a mathematical approach for decision-making in housing planing problems in this paper.

2 Housing planing model by AHP

There are the following three principles when we use the AHP method in problem solving. They are principles of decomposition, comparative judgments, and synthesis of priorities. The decomposition principle requires structuring the hierarchy to capture the basic elements of the problem. An effective way to do this is first to work downward from the focus in the top level to criteria bearing on the focus in the second level, followed by subcriteria in third level, and so on, from the more general to the more particular and definite. The bottom level elements are alternatives. The structure of a hierarchy is given by Figure 1.

The principle of comparative judgments gives the priority of that element which is then used to weight the local priorities of elements in the level below

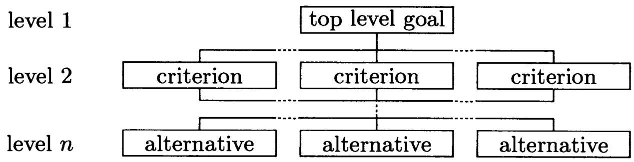


Fig. 1. Structure of a hierarchy

compared by it as criterion, and so on to the bottom level. The measurement standard based on the basic pairwise comparison is given in Table 1.

Table 1. Pairwise comparison

Element i is compared with element j	a_{ij}	a_{ji}
Same important	1	1
A little important	3	1/3
Quite important	5	1/5
Very important	7	1/7
Most important	9	1/9

(2, 4, 6 and 8 also can be used as mean values suitably)

We consider to propose a systematical procedure to select optimal solutions for a user who wants to build his own house. From the point of view of mathematical programming, we introduce a certain type of multicriteria evaluation function associated with a house plan by the following formula

$$\begin{cases} \text{minimize} & (f_1, f_2, f_3) \\ \text{subject to} & (v_1, v_2, v_3), \end{cases}$$

f_1 : construction method, v_1 : economical conditions,
f_2 : room arrangement, v_2 : site-conditions,
f_3 : outside design, v_3 : others.

This research can give a comprehensive evaluation for users' preference with concrete values. At first, a user's plan is quantified by certain evaluations based on each criterion, and then its comprehensive evaluation can be calculated by their weighted sum. If the comprehensive evaluation is not satisfactory for the user, quantifications will be modified by perturbing evaluation values. This evaluation method with multicriteria can be written as follows.

Before building the house, users have to select an optimal construction type from some alternatives of construction styles, such as wooden, 2×4,

Criterion	Quantification	Weight	Comprehensive evaluation
Construction method	f_1	w_1	$\sum_{i=1}^{3} w_i f_i$
Room arrangement	f_2	w_2	
Outside design	f_3	w_3	

↑ modify

Fig. 2. Flow of comprehensive evaluation

and so on. Moreover, outside design is also important for the evaluation as another preference for users; this criterion is treated in [1]. In comparison with these two criteria, room arrangement seems to be more complex and most important for users.

3 Comprehensive evaluation for the house of room arrangements

In this paper, we assume that the land for the house is sufficiently suitable and wide, and that it is satisfactory for general conditions of the single house. Furthermore, we assume that a user is satisfied with this land.

3.1 Calculation of the weight

It can be judged by AHP which item is the most important factor among the three criteria (see [3]). So it is necessary to judge the weight of each item by the information coming from users. It is interpreted as the factor with highest weight being good. Moreover, it turns out that users are thinking which evaluation item as most important by the value of weight. Saaty [2] proposed the AHP absolute measurement method and AHP relative measurement method as algorithm. The alternatives may be evaluated by paired comparisons (relative measurement). Weights for evaluating alternatives are given in Table 2. Similarly, we can obtain weights of other factors by the same method.

Table 2. Pairwise comparison from the single house

	Construction method	Room arrangement	Outside design	Weight
Construction method	1	1/2	4	0.263
Room arrangement	2	1	7	0.658
Outside design	1/4	1/7	1	0.658

3.2 Maximization of utility problem on the floor space

For users as a family, if the area of a site is fixed, there is a problem how the floor space of room arrangement should be chosen. In the same situation, each user may not choose the same plan. In this paper, based on the information came from many users, the utility problem function on selections of the floor space is expressed as follows:

$$U = U(\sum_{i=1}^{n} w_i u_i(x_i)), \quad i = 1, ..., n,$$

where w_i is a weight of selection of floor space x_i, and $u_i(x_i)$ is an evaluation function about selection of floor space x_i. The width of each room is not only restricted to the floor space of a house, but also restricted to the width of the next room. For a user, the biggest concern is to select floor space to the expected maximum utility. We can construct the following function for the optimal floor space which can be determined according to its utility:

$$\begin{cases} \text{minimize } \sum_{i=1}^{n} w_i u_i(x_i) \\ \text{subject to } \sum_{i=1}^{n} x_i \leq L \\ \qquad\qquad x_i \in S_i, \quad i = 1, ..., n, \end{cases}$$

where L is the floor space of a house, and it is related to a loan, and S_i $(i = 1, ..., n)$ is a set of relative to the floor space of each room. Then, we introduce the method of asking for the utility function $u_i(x_i)$. To calculate the coefficient of an utility function by using the certain degree of each floor area. The algorithm for finding the function $u_i(x_i)$ which use Lagrange interpolation polynomial is proposed as follows.

Algorithm

Step(1): The degree f_j is calculated with the weight of each room.

Step(2): The Lagrange interpolation coefficient is construct as follows by using $\pi(x)$:

$$\begin{aligned} I_j(x) &= \pi_n(x_j)/\pi_n(x_n) \\ &= \prod_{\substack{j=1 \\ j \neq n}}^{N} \frac{(x - x_j)}{(x_n - x_j)}. \end{aligned}$$

Step(3): Calculate utility function $u_i(x_i)$ according to the following formula.

$$u_i(x_i) = \sum_{j=1}^{N} I_j(x) f_j.$$

Example

First the weighs of each factor are calculated. The degrees of each factor are calculated possibly if the weights of each factor are known. A degree shows which factor a user thinks as most important. For example, calculate a degree of a kitchen room according to the following Table 3. There values show that the user thinks kitchen room area $9m^2$ is the most impotant. Next the function $u_i(x_i)$ is approximated by using the Lagrange interpolation polynomial by the polynomial which passes though points $\{x_i, f_i\}$. Here x_i and f_i are floor space and degree, respectively. By Table 3, $\{x_1, x_2, x_3, x_4, x_5\} = \{6, 7.5, 9, 10.5, 12\}$ and $\{f_1, f_2, f_3, f_4, f_5\} = \{0.1, 0.7, 1.0, 0.7, 0.4\}$ are floor space and degree of a kitchen room, respectively. Then we can obtain following utility function $u(x)$ of a kitchen room by the upper algorithm,

$$u(x) = 0.007x^4 - 0.259x^3 + 3.250x^2 - 17.117x^1 + 32.200.$$

Similarly, we can obtain others utility functions by the same method.

Table 3. Degree of a kitchen room

	$6m^2$	$7.5m^2$	$9m^2$	$10.5m^2$	$12m^2$	Weight	Degree
$6m^2$	1	1/5	1/7	1/5	1/3	0.043	0.1
$7.5m^2$	5	1	1/3	1	5	0.244	0.7
$9m^2$	7	3	1	1	3	0.336	1.0
$10.5m^2$	5	1	1	1	3	0.257	0.7
$12m^2$	3	1/5	1/3	1/3	1	0.090	0.4

3.3 Comprehensive evaluation system of room arrangements

After the area of each room is decided, the comprehensive evaluation system of room arrangement will be considered. In this paper, a ten-point method is used to evaluate room arrangement. The comprehensive evaluation system is proposed by the following Table 4. The weights of a framework are calculated in Subsection 3.1. The user can obtain the optimal solution easily in an objective framework by using qualitative data, which expresses an evaluation result.

4 Algorithm

House design software can create a floor plan quickly, and we can look at a floor plan from several angles, but it cannot judge the floor plan scientifically. In this paper, users not only plan a wish house by using this software, but also express his preference quantitatively, and evaluate it scientifically. Specifically, the automatic generation algorithm of the house plan arrangement is proposed follows:

Step (1): A user's hope is arranged.
Step (2): After deciding the total floor space, the maximum floor area is calculated. If restriction conditions are satisfied, it will go to the next step, otherwise, it returns to Step (1) and the area of each room is improved.
Step (3): The wish house in Step (1) is designed by using house design software.
Step (4): The house as an image is evaluated. If it satisfies the user's preference sufficiently, the total cost will be calculated and a house maker will be chosen. Otherwise, it returns to Step (1) and improved unreasonable designs.

Table 4. Comprehensive evaluation system

	Item	An evaluating point x	Weight w	Result $\sum wx$
Homely side	number of the rooms	1∼10	0.066	
	width of the room		0.118	=6∼10
	functionality of the room		0.419	(good)
	arrangement of the room		0.111	
	receipt		0.286	=4∼6
Social side	privacy	1∼10	0.392	(good in general)
	Japanese-style room		0.079	general)
	communication space		0.196	
	change of a family		0.333	=1.5∼4
Healthy side	sanitary	1∼10	0.054	(not good)
	sunshine		0.306	
	ventilation		0.154	=1.0∼1.5
	safety		0.486	(quite bad)

5 Conclusion and Remarks

In this paper, a housing planning system is established that can determine several kinds of user's preferences in short time based on a rational decision-making system by using AHP.
It is necessary to determine algorithm which can connect research in this paper with house design software by using the various decision-making technique containing AHP.

References

1. Hui Liang, Construction of supporting system on housing planning based on methods of multicriteria decision making (in Japanese), Master dissertation, Hirosaki University, Hirosaki, 2002.

2. T.L.Saaty: Axiomatic Foundation of the Analytic Hierarchy Process, Management Science 32, 841-855, 1986.
3. Kaoru Tone and Ryutaro Manabe, AHP jireishu (in Japanese), Union of Japanese Scientists and Engineers, 1990.