International Series in
Operations Research & Management Science

Michael Doumpos
Constantin Zopounidis
Evangelos Grigoroudis *Editors*

Robustness Analysis in Decision Aiding, Optimization, and Analytics

Springer

International Series in Operations Research & Management Science

Volume 241

More information about this series at http://www.springer.com/series/6161

Michael Doumpos • Constantin Zopounidis
Evangelos Grigoroudis
Editors

Robustness Analysis in Decision Aiding, Optimization, and Analytics

Springer

Editors

Michael Doumpos
School of Production Engineering
and Management
Financial Engineering Laboratory
Technical University of Crete
Chania, Greece

Evangelos Grigoroudis
School of Production Engineering
and Management
Decision Support Systems Laboratory
Technical University of Crete
Chania, Greece

Constantin Zopounidis
School of Production Engineering
and Management
Financial Engineering Laboratory
Technical University of Crete
Chania, Greece

Audencia Business School
Nantes, France

ISSN 0884-8289 ISSN 2214-7934 (electronic)
International Series in Operations Research & Management Science
ISBN 978-3-319-33119-5 ISBN 978-3-319-33121-8 (eBook)
DOI 10.1007/978-3-319-33121-8

Library of Congress Control Number: 2016943080

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

The Role Robustness in Operations Research and Management Science

Operations research and management science (OR/MS) models are based on assumptions and hypotheses about the available data, the modeling parameters, and the decision context. These are often characterized by uncertainties, fuzziness, vagueness, and errors, which are due to the complexity of real-world problems. As a consequence, it is likely that mild changes on the assumptions and hypotheses set at an early stage of the analysis may require major revisions of the modeling context (e.g., imposing new data requirements, reformulation of objectives, goals, and constraints), thus ultimately leading to very different conclusions and recommendations. Furthermore, it is often observed that solutions found to be acceptable at an early stage of the analysis are actually not easy to implement due to differences (realized a posteriori) between the modeling approach and the actual nature and the evolving dynamic character of the problem at hand.

Robustness analysis seeks to address such issues by promoting models and solutions, which are acceptable under a wide set of plausible conditions and configurations. It is rather difficult to give a unique definition of robustness that fits all contexts and types of problems. However, the common perspective widely used in OR/MS is to consider robustness analysis in the framework of decision-making under uncertainty.

Stewart [8] distinguishes between external and internal uncertainties. External uncertainties relate to the decision environment involving issues that are usually outside the direct control of the decision-maker. Internal uncertainties, on the other hand, relate to problem structuring and modeling issues that arise, for instance, due to the imprecision and ambiguity of judgmental inputs.

Given such uncertainties, Rosenhead [6] highlights the importance of considering the flexibility that solutions/decisions offers. He defines this flexibility as the future opportunity to take decisions toward desired goals. Within this context, he considers the robustness of a solution as the ratio of the number of acceptably

performing configurations with which that solution is compatible to the total number of acceptably performing configurations.

Roy [7], on the other hand, adopts a wider perspective and argues that robustness analysis is a tool that decision analysts use to protect against the *approximations* and *ignorance zones*, which arise due to imperfect knowledge, ill-defined data, and the specification of modeling parameters. Such issues create a gap between the "true" model and the one resulting from a computational mechanism. Roy views the characterization of robustness solely in the context of uncertainty as a restrictive approach and suggests instead going beyond the traditional scenario-based approach through the adoption of a version/procedure-based framework that takes into account different realities for a problem (versions) and processing procedures. This is similar to the approach proposed by Vincke [9] who described robust solutions as those that remain acceptable under changes in the problem data and the parameters of the method used while further highlighting that robustness also applies to the decision methods used to derive the results of an analysis.

Similar views can also be found in the context of robust optimization, which has been an active research topic in OR/MS at least since the 1990s [1–3]. For instance, Mulvey et al. [5] distinguish between the robustness of solutions for a given problem which are acceptable under different modeling forms and the robustness of the modeling scheme. They note that *reactive* approaches relying on post-optimality techniques (e.g., sensitivity analysis) are not enough as they only take into account data uncertainties, thus proposing the use of *proactive* approaches, which focus on formulations that, by design, provide less sensitive (more robust) solutions to changes in the problem data. Mulvey et al. further distinguish the robust optimization paradigm from traditional OR/MS approaches such as stochastic programming. The differences between these approaches are also analyzed by Kouvelis and Yu [4] who provide a formal framework for robust optimization with emphasis on discrete optimization problems.

All the above different views of robustness cover a broad OR/MS context that starts from soft OR and decision-aiding tools and extends to a wide range of analytical techniques for different types of optimization problems. As new challenges emerge in a "big-data" era, where the information volume, speed of flow, and complexity increase rapidly, and analytics playing a fundamental role for strategic and operational decision-making at a global level, robustness issues such as the ones outlined above become more relevant than ever for providing sound decision support through more powerful analytic tools.

Outline of the Book

Aims and Scope

Given the multifaceted nature of robustness, the motivation for the preparation of this book was to publish a unique volume aiming at providing a broad coverage

of the recent advances in robustness analysis in decision aiding, optimization, and analytics, adopting an OR/MS perspective.

The board coverage of the volume is a unique feature that enables the comprehensive illustration of the challenges that robustness raises in different OR/MS contexts and the methodologies proposed from multiple perspectives. Thus, this edited volume facilitates the presentation of the current state of the art and the communication of ideas, concepts, and techniques for different OR/MS areas where robustness concerns are highly relevant.

The volume also includes a part on applications of robust techniques in engineering and management, thus illustrating the robustness issues raised in real-world problems and their resolution with the lasted advances in robust analytical techniques.

Organization

The book includes 14 chapters, organized in three main parts that cover a wide range of topics related to theoretical advances in robustness analysis and their applications. The first part is devoted to decision aiding. The book starts with the chapter of Lahdelma and Salminen about stochastic multicriteria acceptability analysis (SMAA). SMAA is a popular approach for multicriteria decision aid (MCDA) problems under uncertainty. SMAA enables the evaluation of a discrete set of decision alternatives when there is uncertainty about the data and/or the parameters of the decision model. Uncertainty is represented through probability distributions, and probabilistic indicators are constructed that facilitate the formulation of robust recommendations. The chapter illustrates the main concepts and functionality of this approach using an easy-to-follow example-based illustration. Implementation issues and recent advances are further discussed.

The second chapter, by Doumpos and Zopounidis, focuses on preference disaggregation analysis (PDA). PDA is widely used in MCDA to infer decision models from data using optimization-based techniques (usually linear programming models). Over the past decade, much research has been devoted on the development of robust PDA approaches that take into consideration a set of decision models (of the same type/class) rather than a single model. The chapter examines the robustness of such approaches in classification problems, where a finite set of alternative should be classified into predefined performance categories. The chapter proposes new robustness indicators based on concepts and techniques from the field of convex optimization, taking into account the geometric properties of the set of feasible/acceptable values for the parameters of a decision model as specified by a set of decision instances. The new indicators are illustrated and validated through a numerical example.

The third chapter of this first part of the book, by Ríos Insua, Ruggeri, Alfaro, and Gomez, is devoted to adversarial risk analysis (ARA), which is a risk management framework for decision situations involving intelligent opponents. ARA has been

recently applied in a wide range of areas, including business applications, defense, and security. The latter is the main focus of the chapter, which provides an outline of the role of robust methods in ARA. The chapter starts by discussing Bayesian robustness and then presents a game theoretic framework applied to sequential and simultaneous defend-attack instances. The framework leads to game theoretic solutions, which are improved through robustness analysis and ARA.

The first part of the book closes with the chapter by Sniedovich about Wald's maximin paradigm, which has played a central role in decision-making under uncertainty, as a tool for worst-case-based robustness analysis. The chapter presents the conceptual and modeling aspects of the Wald's maximin paradigm and analyzes its differences from other similar frameworks. The relationship between this paradigm and robust decision-making is also discussed, from the perspective of robust optimization, where the maximin principle has been extensively used for coping with different types of robustness issues.

The second part of the book contains four chapters about robust optimization. This part starts with the overview paper of Sözüer and Thiele. The authors provide a survey of the most recent advances in the theory and applications of robust optimization over the past 5 years (2011–2015). The survey covers methodological issues related to static and multistage decision-making, stochastic optimization, distributional robustness, and nonlinear optimization, as well as a range of application areas such as supply chain management, finance, revenue management, and health care.

In the next chapter, Kasperski and Zieliński focus on robustness for discrete optimization problems and discusses the two most popular approaches of modeling the uncertainty, namely, the discrete and interval uncertainty representations. The chapter starts with describing the traditional minimax approach and proceeds with the presentation of new concepts and techniques that have recently appeared in the literature, such as the use of weighted ordering averaging, robust optimization with incremental recourse, and two-stage problems. Computational complexity issues, which are very relevant for this type of problems, are also discussed.

The third chapter in this part, by Chassein and Goerigk, discusses the assessment of robust solutions in optimization problems. This is a relevant issue, given the wide range of definitions of robustness concepts, criteria, and metrics, available in the literature, which naturally create a confusion regarding the selection of the most appropriate approach for a given problem. The chapter illustrates this issue using as examples well-known optimization problems, namely, the assignment and knapsack problems, and proposes formal evaluation frameworks. These are illustrated through experimental data.

In the last chapter of the second part, Inuiguchi examines fuzzy linear programming (LP) problems. Fuzzy optimization enables the modeling of decision problems that incorporate ambiguity and vagueness. This chapter focuses on LPs with fuzzy coefficients in the objective functions. Robustness analysis in this context is more involved compared to traditional optimization problems. Inuiguchi defines two approaches based on the minimax and maximin principles. Algorithmic and computational issues that arise in the implementation of the proposed approaches are also analyzed.

The last part of the book is devoted to application of robust OR/MS techniques in engineering and management. This part includes six chapters. The first of these chapters, by Artigues, Billaut, Cheref, Mebarki, and Yahouni, considers robust machine scheduling problems under uncertainty with a group sequence structure, where an ordered partition of jobs is assigned to each machine. Standard robust scheduling techniques are reviewed together with recoverable robust optimization methods. Empirical evidence derived from a real manufacturing system is also reported.

The next two chapters involve applications related to policy decision-making for environmental management and energy systems. In particular, Kwakkel, Eker, and Pruyt adopt a multi-objective optimization framework. The authors consider a case study related to the European policies for reducing carbon emissions and promoting the use of renewable energy technologies. A system dynamics model is used to simulate paths for the European electricity system, considering a number of uncertain inputs variables. The policy design problems is formulated as an optimization problem with three objectives, and different robustness metrics are applied to examine which is the most appropriate one for the making robust policy recommendations.

The next chapter, by Nikas and Doukas, presents a framework based on fuzzy cognitive mapping for selecting effective climate policies for low carbon transitions in the European Union. The proposed approach is an analytical framework for developing robust transition pathways, grounded on existing quantitative models, an extensive literature review of the risks and uncertainties involved, and qualitative information deriving from a structured stakeholder engagement process.

The next two chapters focus on portfolio optimization. The uncertainties that prevail in the financial markets have attracted a lot of interest for robust techniques in this area. The chapter of Gülpınar and Hu presents an overview of the theory and applications of robust approaches to portfolio optimization, focusing on the most fundamental and widely studied single-period context. The authors discuss the relevance of using symmetric and asymmetric uncertainty sets for modeling asset returns, cover recent advances in recent developments in data-driven robust optimization, and discuss the connections between robust optimization and financial risk management.

In the next chapter, Keçeci, Kuzmenko, and Uryasev consider portfolio optimization with stochastic dominance constraints. Stochastic dominance provides a distribution-free approach that takes into account the entire returns' distribution. The authors present efficient numerical algorithms for solving optimization problems with second-order stochastic dominance constraints. Empirical results are presented based on data from the Dow Jones and DAX indices in comparison to the well-known mean-variance portfolio optimization model.

The book closes with the chapter of Atıcı and Gülpınar about performance and production efficiency measurement, in the context of data envelopment analysis (DEA). DEA is widely used as a nonparametric efficiency assessment technique, based on linear programming models. In this chapter, the authors consider the DEA framework under uncertainty about the data (input/outputs). An imprecise DEA approach and a robust optimization model are compared using a case study involving

the assessment of production efficiency from the agricultural sector (olive-growing farms). The results lead to insights about how the treatment of uncertainty relates to the obtained efficiency estimates.

Acknowledgments

Closing this brief editorial, we should express our sincere thanks and gratitude to all authors who accepted our invitation to contribute to this project, devoting considerable time and effort to prepare excellent comprehensive works of high scientific quality and value. Without their help, it would be impossible to prepare this book in line with the high standards that we have set from the very beginning of this project.

Chania, Greece

Michael Doumpos
Constantin Zopounidis
Evangelos Grigoroudis

References

1. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. Math. Oper. Res. **23**(4), 769–805 (1998)
2. Ben-Tal, A., Nemirovski, A.: Robust solutions of uncertain linear programs. Oper. Res. Lett. **25**(1), 1–13 (1999)
3. Bertsimas, D., Brown, D.B., Caramanis, C.: Theory and applications of robust optimization. SIAM Rev. **53**(3), 464–501 (2011)
4. Kouvelis, P., Yu, G.: Robust Discrete Optimization and its Applications. Kluwer Academic Publishers, Boston (1997)
5. Mulvey, J.M., Vanderbei, R.J., Zenios, S.A.: Robust optimization of large-scale systems. Oper. Res. **43**(2), 264–281 (1995)
6. Rosenhead, J.: Robustness analysis: keeping your options open. In: Rosenhead, J., Mingers, J. (eds.) Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflicts, pp. 181–207. Wiley, Chichester (2001)
7. Roy, B.: Robustness in operational research and decision aiding: a multi-faceted issue. Eur. J. Oper. Res. **200**(3), 629–638 (2010)
8. Stewart, T.: Dealing with uncertainties in MCDA. In: Figueira, J., Greco, S., Ehrogott, M. (eds.) Multiple Criteria Decision Analysis: State of the Art Surveys, pp. 445–466. Springer, New York (2005)
9. Vincke, P.: Robust solutions and methods in decision-aid. J. Multicrit. Decis. Anal. **8**(3), 181 (1999)

Contents

List of Contributors

Cesar Alfaro
Department of Statistics and Operations Research, Rey Juan Carlos University, Fuenlabrada, Madrid, Spain

Christian Artigues
CNRS, LAAS, Toulouse, France
Univ de Toulouse, LAAS, Toulouse, France

Kazım Barış Atıcı
Department of Business Administration, Hacettepe University, Ankara, Turkey

Jean-Charles Billaut
Laboratoire d'Informatique de l'Université de Tours, Tours, France

André Chassein
University of Kaiserslautern, Kaiserslautern, Germany

Azzedine Cheref
CNRS, LAAS, Toulouse, France
Univ de Toulouse, LAAS, Toulouse, France

Haris Doukas
School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece

Michael Doumpos
School of Production Engineering and Management, Financial Engineering Laboratory, Technical University of Crete, Chania, Greece

Sibel Eker
Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

Neslihan Fidan Keçeci
School of Business, Istanbul University, Istanbul, Turkey

Marc Goerigk
Lancaster University, Lancaster, UK

Javier Gomez
Department of Statistics and Operations Research, Rey Juan Carlos University, Fuenlabrada, Madrid, Spain

Nalân Gülpınar
Warwick Business School, The University of Warwick, Coventry, UK

Zhezhi Hu
Warwick Business School, The University of Warwick, Coventry, UK

Masahiro Inuiguchi
Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka, Japan

Adam Kasperski
Faculty of Computer Science and Management, Wrocław University of Technology, Wrocław, Poland

Viktor Kuzmenko
Glushkov Institute of Cybernetics, Kyiv, Ukraine

Jan H. Kwakkel
Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

Risto Lahdelma
Energy Technology, Aalto University, Otakaari 4, FIN-02015 Espoo, Finland

Nasser Mebarki
LUNAM, Université de Nantes, IRCCyN Institut de Recherche en Communication et Cybernétique de Nantes, UMR CNRS, Nantes, France

Alexandros Nikas
School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece

Erik Pruyt
Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

David Ríos Insua
Instituto de Ciencias Matemáticas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

Fabrizio Ruggeri
Consiglio Nazionale delle Ricerche, Istituto di Matematica Applicata e Tecnologie Informatiche, Milano, Italy

Pekka Salminen
School of Business and Economics, University of Jyväskylä, Jyväskylä, Finland

Moshe Sniedovich
Department of Mathematics and Statistics, University of Melbourne, Melbourne, VIC 3010, Australia

Seçil Sözüer
Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA, USA

Aurélie C. Thiele
Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA, USA

Stan Uryasev
Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA

Zakaria Yahouni
LUNAM, IRCCyN/MELT, Manufacturing Engineering Laboratory of Tlemcen, Tlemcen University, Tlemcen, Algeria

Paweł Zieliński
Faculty of Fundamental Problems of Technology, Wrocław University of Technology, Wrocław, Poland

Constantin Zopounidis
School of Production Engineering and Management, Financial Engineering Laboratory, Technical University of Crete, Chania, Greece
Audencia Business School, Nantes, France

Chapter 1
SMAA in Robustness Analysis

Risto Lahdelma and Pekka Salminen

Abstract Stochastic multicriteria acceptability analysis (SMAA) is a simulation based method for discrete multicriteria decision aiding problems where information is uncertain, imprecise, or partially missing. In SMAA, different kind of uncertain information is represented by probability distributions. Because SMAA considers simultaneously the uncertainty in all parameters, it is particularly useful for robustness analysis. Depending on the problem setting, SMAA determines all possible rankings or classifications for the alternatives, and quantifies the possible results in terms of probabilities. This chapter describes SMAA in robustness analysis using a real-life decision problem as an example. Basic robustness analysis is demonstrated with respect to uncertainty in criteria and preference measurements. Then the analysis is extended to consider also the structure of the decision model.

1.1 Introduction

Robustness analysis of a computational model is a type of sensitivity analysis that considers simultaneous variations of all parameters in a given domain. More general robustness analysis would also consider the sensitivity of the analysis with respect to model structure derived from various assumptions. Robustness analysis is necessary in particular when some input parameters of the model are imprecise or uncertain.

R. Lahdelma (✉)
Energy Technology, Aalto University, Otakaari 4, FIN-02015 Espoo, Finland
e-mail: risto.lahdelma@aalto.fi

P. Salminen
School of Business and Economics, University of Jyväskylä, P.O. Box 35,
FIN-40014 Jyväskylä, Finland
e-mail: pekka.o.salminen@jyu.fi

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_1

Stochastic multicriteria acceptability analysis (SMAA) is a simulation based method for discrete multicriteria decision aiding problems where information is uncertain, imprecise, or partially missing. In SMAA, different kind of uncertain information is represented by probability distributions. This approach is similar to metrology [22]. For example, if the cost of an alternative is not accurately known, it can be represented by a uniform distribution in a given range, or a normal distribution with specified expected value and standard deviation (Fig. 1.1). Uncertain preference information is similarly represented by distributions. Also subsequent computations in SMAA follow probability theory.

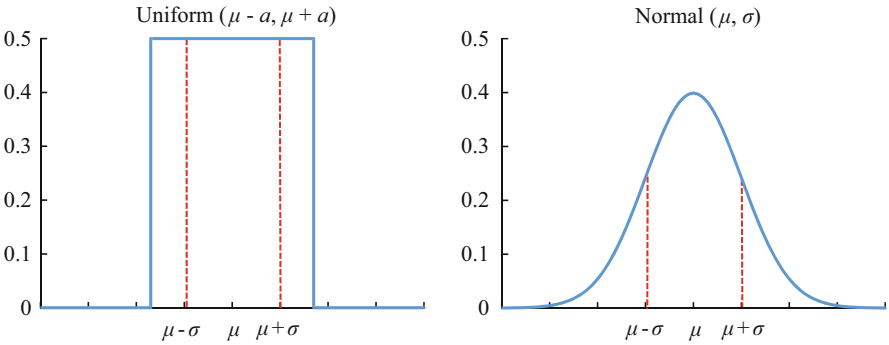


Fig. 1.1: Representing uncertain criteria measurements as distributions

Depending on the problem setting, SMAA computes statistically for each alternative the probability to be most preferred, dominate another alternative, be placed on a particular rank or fit in a specific category. The computation is implemented by Monte-Carlo simulation, where values for the uncertain variables are sampled from their distributions and alternatives are evaluated by applying the decision model.

SMAA can be applied with different decision models. These include linear and non-linear utility or value functions [8, 15, 16], ELECTRE methods [9, 26], reference point based methods [11, 17], efficiency score of Data Envelopment Analysis (DEA) [10], nominal classification method [29], and ordinal classification method [12]. For a surveys on different variants and applications of SMAA, see [13, 24]. Recent developments of SMAA include robustness analysis with respect to shape of the utility function by Lahdelma and Salminen [14], efficient Markov Chain Monte Carlo simulation technique to treat complex preference information by Tervonen et al. [27], the SMAA-PROMETHEE method by Corrente et al. [4], SMAA with Choquet Integral by Angilella et al. [1], and extensions for pairwise comparison methods such as the analytic hierarchy process (AHP) by Durbach et al. [6] and the Complementary Judgment Matrix (CJM) method by Wang et al. [28].

Because SMAA considers simultaneously the uncertainty in all parameters, it is particularly useful for robustness analysis of different multicriteria decision models. SMAA determines all possible rankings or classifications for the alternatives, and quantifies the possible results in terms of probabilities. The solution with highest

probability is typically the recommended solution. However, the probabilities for other possible solutions are also provided for the decision makers (DMs). This means that SMAA describes how robust the model is subject to different uncertainties in the input data. SMAA can also be used to analyze the robustness of the decision problem with respect to the model structure. For example, robustness with respect to linearity assumptions in utility/value functions can be analyzed by choosing a more general parametrized utility function and exploring how the solutions change as a function of the degree of non-linearity.

In the following, we describe the SMAA method applied on a real-life decision problem of choosing a location for an air cargo hub in Morocco [21]. Section 1.2 describes problem representation in SMAA as a stochastic MCDA problem and how it is analysed using stochastic simulation. Section 1.3 presents the statistical measures of SMAA and shows how SMAA can be used to assess the robustness of an MCDA problem with respect to uncertainty in criteria and preference measurements. Section 1.4 extends the robustness analysis to consider the structure of the decision model.

1.2 Problem Representation in SMAA

1.2.1 Stochastic MCDA Problem

A discrete multi-criteria decision problem consists of a set of m alternatives that are measured in terms of n criteria. The alternatives are evaluated using a decision model $M(\mathbf{x}, \mathbf{w})$ that depends on the applied decision support method. The matrix $\mathbf{x} = [x_{ij}]$ contains the criteria measurements for each alternative i and criterion j. The preference information vector $\mathbf{w} = [w_j]$ represents the DM's preferences. Typically $\mathbf{w}$ contains importance weights for the criteria. Depending on the decision model, $\mathbf{w}$ can also contain other preference parameters, such as various shape parameters for non-linear models.

SMAA has been developed for real-life problems, where both criteria and preference information can be imprecise, uncertain or partially missing. To represent the incompleteness of the information explicitly, SMAA represents the problem as a *stochastic MCDA model*, where criteria and preference information is represented by suitable (joint) probability distributions:

- $f_X(\mathbf{x})$ the density function for stochastic criteria measurements.
- $f_W(\mathbf{w})$ the density function for stochastic importance weights or other preference parameters.

Because all information is represented uniformly as distributions, this allows using efficient simulation techniques for analyzing the problem and deriving results about prospective solutions and their robustness.

An example of a stochastic MCDA model is the problem of choosing a location for a centralized air cargo hub in Morocco [21]. In this problem, nine alternative

locations were considered. Different socio-economic factors, the geographical location, and environmental impacts were formalized as six criteria: INVEST = investment cost, PROXIMITY = proximity to producers, POTENTIAL = potential of the site, TRANSPORT = transport cost, SERVICE = service level, ENVIRON = Environment. The alternatives, criteria and measurements are presented in Table 1.1.

The INVEST, POTENTIAL, TRANSPORT and SERVICE criteria were measured on cardinal scales. The values in Table 1.1 for these criteria are dimensionless quantities that have been obtained by scaling the actual measurements on linear scales where larger values are better. The uncertainty of these measurements appears on the last row as a plus/minus percentage. The measurements were then represented as independent, uniformly distributed random numbers in the plus/minus ranges around their expected values. In SMAA it is possible to use arbitrary distributions to represent uncertain criteria measurements. If the uncertainties of the criteria measurements are dependent, this can be represented by joint distributions, such as the multivariate Gaussian distribution [18, 19].

The PROXIMITY and ENVIRON criteria were evaluated ordinally, i.e. experts ranked the alternatives with respect to these criteria so that the best alternative obtained rank 1, second best rank 2 etc. Ordinal measurement can be necessary if cardinal measurement is too costly, or if it is difficult to form a measurable scale for the criterion.

Table 1.1: Alternatives and criteria measurements in air cargo hub case (alphabetical order)

Alt	INVEST (max)	PROXIMITY (min)	POTENTIAL (max)	TRANSPORT (max)	SERVICE (max)	ENVIRON (min)
Agadir	70	2	165	644	50	2
Benslimane	80	3	560	3718	40	1
Casablanca	65	1	585	3621	80	5
Dakhla	80	8	82	600	20	1
Fez	70	6	385	2872	30	4
Marrakesh	65	5	379	2589	45	1
Oujda	75	7	82	663	25	4
Rabat	65	4	542	3718	45	3
Tangier	70	3	357	1915	60	3
Uncertainty	±10 %	Ordinal	±10 %	±10 %	±10 %	Ordinal

1.2.2 Generic SMAA Simulation

Different variants of SMAA apply the generic simulation scheme of Algorithm 1 for analyzing stochastic MCDA problems. During each iteration, criteria measurements, weights, and possible other preference parameters are drawn from their dis-

tributions, and the decision model is used to evaluate the alternatives. Depending on the problem setting and decision model, different statistics are collected during the simulation and the SMAA measures are computed based on the statistics. For example, in the case of a ranking problem, statistics are collected about how frequently alternatives obtain a given rank.

Algorithm 1. Generic SMAA simulation

Assume a decision model $M(\mathbf{x}, \mathbf{w})$ for ranking or classifying the alternatives using precise information (criteria matrix $\mathbf{x}$ and preference information vector $\mathbf{w}$)

Use Monte-Carlo simulation to treat stochastic criteria and preference parameters:
Repeat Ktimes {
Draw $<\mathbf{x}, \mathbf{w}>$ from their distributions
Rank, sort or classify the alternatives using $M(\mathbf{x},\mathbf{w})$
Update statistics about alternatives
}

Compute results based on the collected statistics

The efficient implementation and computational efficiency of SMAA methods have been described by Tervonen and Lahdelma [25]. The computational accuracy of the main results depends on the square root of the number of iterations, i.e. increasing the number of iterations by a factor of 100 will increase the accuracy by one decimal place. In practice about 10,000 iterations yield sufficient accuracy for the SMAA results.

1.2.3 Decision Model

SMAA can be used with arbitrarily shaped utility functions, and also with other kinds of decision models that are based on any kind of preference parameters. A common type of utility function is the *additive* utility function that defines the overall utility as a weighted sum of *partial utilities*:

$$u(x_i, \mathbf{w}) = w_1 u_{i1} + w_2 u_{i2} + \ldots + w_n u_{in} \tag{1.1}$$

The w_j are the importance weights for criteria and u_{ij} are the partial utilities obtained by mapping the original criteria measurements (expressed in various units) to unit-less scales so that the worst outcome is 0 and the best outcome is 1. The mappings can be linear or non-linear monotonic functions.

In the sample problem linear mappings were applied, leading to a linear overall utility function. In this study we consider also non-linear mappings in order to analyze the robustness of the problem with respect to the shape of the utility function.

The weights should be non-negative and normalized so that their sum is 1. By substituting 1 or 0 for each partial utility in (1.1) we see that the overall utility is 1 for an ideal alternative, and 0 for an anti-ideal alternative.

1.2.4 Preference Information

In SMAA, incomplete preference information is represented using probability distributions. In the following we consider incomplete weight information. However the same techniques can be used also for other preference parameters.

With an additive utility function, the weights express the relative importance of raising each criterion from its worst value to the best value. Ratios between weights correspond to trade-offs between criteria. In SMAA uncertain or imprecise weights are represented as a joint probability distribution in the *feasible weight space* defined as the set of non-negative and normalized weights

$$W = \{\mathbf{w} \,|\, w_j \geq 0 \text{ and } w_1 + w_2 + \ldots + w_n = 1\} \tag{1.2}$$

This means that the feasible weight space is an $(n-1)$-dimensional simplex. In the 3-criterion case, the feasible weight space is a triangle with corners (1,0,0), (0,1,0) and (0,0,1), as illustrated in Fig. 1.2a. In the absence of weight information, we assume that any feasible weights are equally possible, which is represented by a uniform distribution in W.

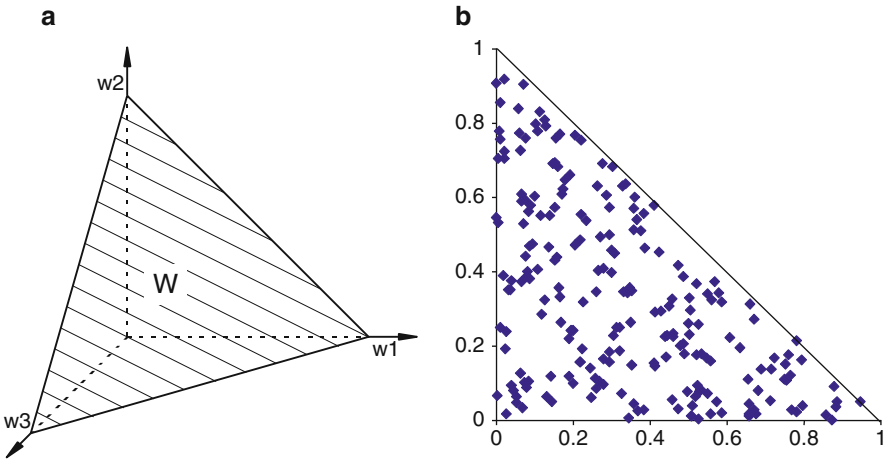


Fig. 1.2: (**a**) Feasible weight space in the 3-criterion case. (**b**) Sampling uniformly distributed weights in the 3-criterion case projected on the (w_1, w_2) plane

Uniformly distributed normalized weights need to be generated using a special technique [25]. First $n-1$ independent uniformly distributed random numbers in the interval [0,1] are generated and sorted together with 0 and 1 into ascending order to get $0 = r_0 \leq r_1 \leq \cdots \leq r_n = 1$. From these numbers the weights are computed as the intervals $w_1 = r_1 - r_0$, $w_2 = r_2 - r_1$, $\ldots$, $w_n = r_n - r_{n-1}$. It is obvious that the resulting weights will be non-negative and normalized. For the proof that the resulting joint distribution is uniform, see [5]. Figure 1.2b illustrates generation of uniformly distributed weights in the 3-dimensional case, projected on the (w_1, w_2) plane where $w_3 = 1 - w_1 - w_2$.

Preference information can be treated in SMAA by restricting the uniform weight distribution with additional constraints. Another technique is to apply a non-uniform distribution for the weights. For example, if the DMs express precise weights with implicit imprecision, this can be represented by a distribution with decreasing density around the expressed weights. Suitable distributions are e.g. triangular distributions and (truncated) normal distributions.

Different ways to restrict the uniform or non-uniform weight distribution with additional constraints include the following:

- Weight intervals can be expressed as $w_j \in [w_j^{\min}, w_j^{\max}]$. Weight intervals may result from DMs' preference statements of type "the importance weight for criterion j is between $w_j^{\min}$ and $w_j^{\max}$". Weight intervals can also be computed to include precise weights or weight intervals of a group of DMs. Figure 1.3a illustrates weight intervals in the 3-criterion case.
- Intervals for trade-off ratios between criteria can be expressed as $w_j/w_k \in [w_{jk}^{\min}, w_{jk}^{\max}]$. Such intervals may result from preference statements like "criterion j is from $w_{jk}^{\min}$ to $w_{jk}^{\max}$ times more important than criterion k". These intervals can also be determined to include the preferences of a group of DMs. Figure 1.3b illustrates two constraints for trade-off ratios.
- Ordinal preference information can be expressed as linear constraints $w_1 \geq w_2 \geq \cdots \geq w_n$. Such constraints represent DMs preference statement that the criterion 1 is most important, 2 is second etc. It is also possible to allow unspecified importance ranking for some criteria or equal importance ($w_j = w_k$). Multiple DMs may either agree on a common partial ranking, or they can provide their own rankings, which can then be combined into a partial ranking that is consistent with each DM's preferences. Figure 1.3c illustrates ordinal preference information.
- DMs holistic preference statements "alternative x_i is more preferred than x_k" result in constraints $u(x_i, \mathbf{w}) \geq u(x_k, \mathbf{w})$ for the weights. In the case of an additive utility/value function, these constraints will be linear inequalities in the weight space. Figure 1.3d illustrates one such holistic preference statement. In the general case, with non-additive utility/value functions, outranking models etc., holistic constraints correspond to non-linear constraints in the weight space.

Weight constraints can be implemented by modifying the weight generation procedure to reject weights that do not satisfy the constraints. In most cases this technique is very efficient. In some cases the Markov Chain Monte Carlo simulation technique is more efficient [27].

1.2.5 Cardinal Criteria

In the case of a linear utility function, the partial utilities u_{ij} are computed from the actual cardinal criteria measurements x_{ij} through linear scaling. The best and worst values can be determined as some natural ideal and anti-ideal values, if such

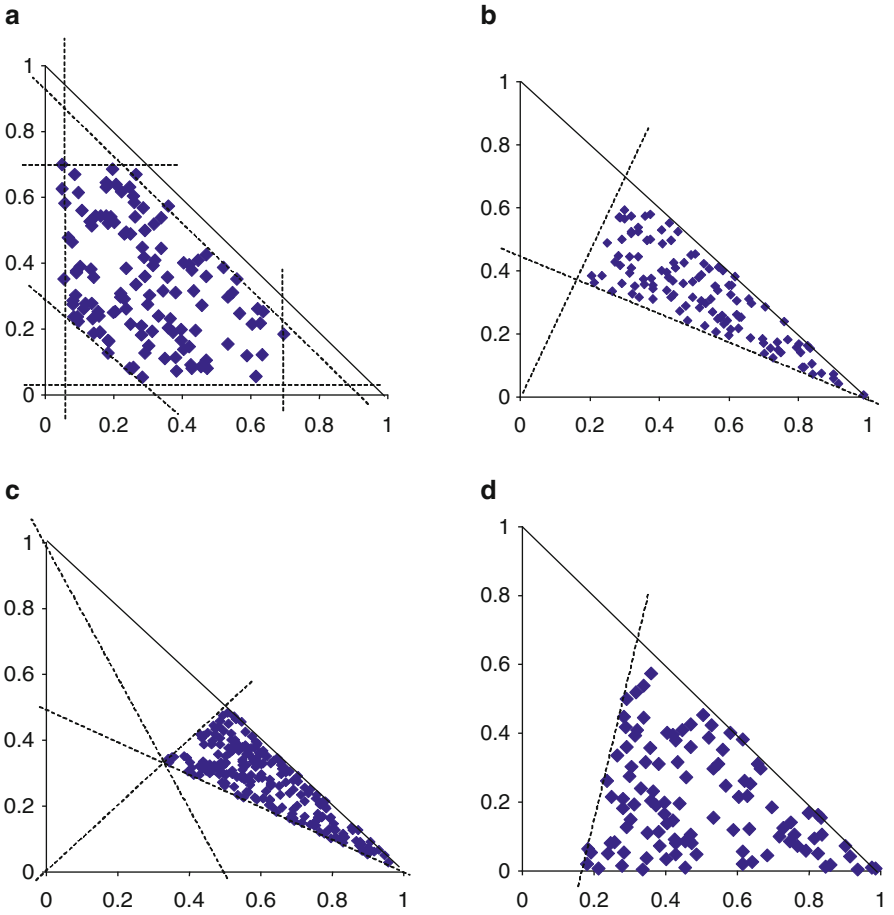


Fig. 1.3: Sampling uniformly distributed weights in the 3-criterion case projected on the (w_1, w_2) plane: (**a**) with interval constraints for weights; (**b**) with two constraints for trade-off ratios; (**c**) with ordinal preference information $w_1 \geq w_2 \geq w_3$; (**d**) with holistic preference information based on an additive utility/value function

exist. For example, the ideal value for costs could be 0 and the ideal value for an efficiency ratio could be 100 %. If such ideal and anti-ideal values cannot easily be defined, it is possible to do the scaling according to the best and worst measurements among the alternatives, as has been done for the sample problem in Table 1.2. Also the uncertainties have been scaled accordingly. A downside with scaling based on best and worst criteria measurements is that the scaling may change if the set of alternatives or their measurements change during the decision process.

As a result, the uncertainty intervals may contain values outside the [0, 1] range. This is not a problem, because the scaling interval is arbitrary; any other interval would order the alternatives identically according to their utilities.

Table 1.2: Scaled cardinal criteria measurements and their uncertainties in air cargo hub case

Alt	INVEST	POTENTIAL	TRANSPORT	SERVICE
Agadir	0.33±0.47	0.17±0.03	0.01±0.02	0.50±0.08
Benslimane	1.00±0.53	0.95±0.11	1.00±0.12	0.33±0.07
Casablanca	0.00±0.43	1.00±0.12	0.97±0.12	1.00±0.13
Dakhla	1.00±0.53	0.00±0.02	0.00±0.02	0.00±0.03
Fez	0.33±0.47	0.60±0.08	0.73±0.09	0.17±0.05
Marrakesh	0.00±0.43	0.59±0.08	0.64±0.08	0.42±0.08
Oujda	0.67±0.50	0.00±0.02	0.02±0.02	0.08±0.04
Rabat	0.00±0.43	0.91±0.11	1.00±0.12	0.42±0.08
Tangier	0.33±0.47	0.55±0.07	0.42±0.06	0.67±0.10

1.2.6 Ordinal Criteria

Ordinal criteria measurements are imprecise: we know the rank of each alternative with respect to the ordinal criterion, but we do not know how much better the first alternative is than the second or third one, etc. In SMAA, ordinal criteria are treated by simulating cardinal values that are consistent with the given ordinal ranks. The first rank corresponds to cardinal value $s_1 = 1$ and the last rank R corresponds to $s_R = 0$. The intermediate ranks 2, 3, …, $R-1$ should correspond to a descending sequence of unknown cardinal values between 1 and 0. To obtain the unknown intermediate values, $R-2$ independent uniformly distributed random numbers in the interval [0, 1] are generated. These values are then sorted together with 1 and 0 into descending order to obtain cardinal values that satisfy $1 = s_1 \geq s_2 \geq \cdots \geq s_{R-1} \geq s_R = 0$.

The process described converts ordinal criteria into stochastic cardinal criteria. Note that the intervals between subsequent values $s_r - s_{r+1}$ are non-negative and their sum is 1. Subject to these constraints, the intervals follow a uniform distribution [5].

In the air cargo hub case, the PROXIMITY and ENVIRON criteria were ordinal. Figure 1.4 shows some random cardinal mappings for these criteria. For the PROXIMITY criteria, alternatives Benslimane and Tangier were both ranked on level 3. Therefore rank levels 1–8 were assigned for the nine alternatives. Similarly, shared ranks for the ENVIRON criteria resulted in assigning five different rank levels for that criterion.

1.3 Robustness with Imprecise Criteria and Weights

In the following we demonstrate the SMAA method using the air cargo hub case presented in Sect. 1.2. A linear utility/value function was used as the decision model in this application. The simulation scheme presented in Algorithm 1 is applied and the utility function is used to rank the alternatives. Observe that this approach

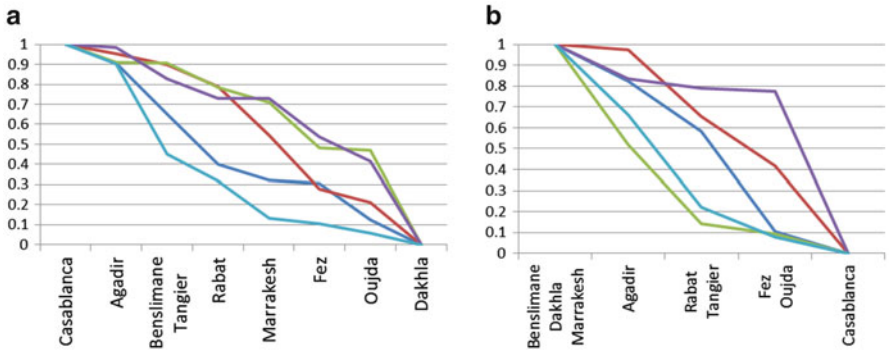


Fig. 1.4: Sample of simulated cardinal values for the air cargo hub case. (**a**) PROXIMITY criterion. (**b**) ENVIRON criterion

differs from traditional utility function methods that compute the expected utility. This means that SMAA does not require a cardinal utility function—an ordinal utility/value function is sufficient. Based on the ranking, the following statistics are collected during the simulation:

- B_{ir}: The number of times alternative x_i obtained rank r.
- C_{ik}: The number of times alternative x_i was more preferred than x_k.
- $\mathbf{W}_i$: Sum of the weight vectors that made alternative x_i most preferred.

Based on the collected statistics the basic SMAA measures are computed. These include *rank acceptability indices*, *pairwise winning indices*, *central weight vectors*, and *confidence factors*, as presented in the following sections.

1.3.1 Rank Acceptability Indices

The primary SMAA measure is the *rank acceptability index* b_i^r. It measures the variety of different preferences that place alternative x_i on rank r. It is the share of all feasible weights that make the alternative acceptable for a particular rank. In other words, it is the probability that the alternative obtains a certain rank. Particularly interesting is the first rank acceptability index b_i^1, which is the probability that the alternative is the most preferred one. For inefficient alternatives the first rank acceptability index is zero. The rank acceptability indices are estimated from the simulation statistics (with K iterations) as

$$b_i^k \approx B_{ir}/K \tag{1.3}$$

The rank acceptability indices can be used for ***robust choice of one or a few best alternatives*** from a large set. Alternatives with high acceptability for the best ranks are candidates for the most acceptable solution. Alternatives with large acceptability

Table 1.3: Rank acceptability indices for air cargo hub case (sorted by b_1)

Alt	b^1	b^2	b^3	b^4	b^5	b^6	b^7	b^8	b^9
Benslimane	72.00	23.00	4.00	1.00	0.00	0.00	0.00	0.00	0.00
Casablanca	25.00	33.00	14.00	7.00	5.00	5.00	5.00	3.00	3.00
Dakhla	1.00	7.00	6.00	5.00	5.00	6.00	15.00	41.00	15.00
Agadir	0.40	5.00	9.00	11.00	13.00	20.00	30.00	9.00	3.00
Tangier	0.38	7.00	17.00	28.00	25.00	16.00	4.00	2.00	0.00
Rabat	0.28	14.00	37.00	17.00	13.00	9.00	5.00	3.00	2.00
Marrakesh	0.03	11.00	11.00	22.00	25.00	19.00	6.00	3.00	3.00
Oujda	0.02	0.00	1.00	2.00	2.00	2.00	4.00	20.00	70.00
Fez	0.00	1.00	2.00	7.00	12.00	23.00	31.00	19.00	4.00

for the worst ranks should be avoided when searching for a robust most preferred alternative even if they would have fairly high acceptability for the best ranks. If none of the alternatives receive high acceptability indices for the best ranks, it indicates a need to measure the criteria, preferences or both more accurately.

Table 1.3 presents the rank acceptability indices for the air cargo hub case and Fig. 1.5 shows the corresponding *acceptability profile*. To make the acceptability profile easy to read, the alternatives are sorted by their first rank acceptability index. In case of equal first rank indices, order is determined based on the second index etc. This is called *lexicographic order*. The most acceptable (best) alternatives are Benslimane and Casablanca with clearly highest acceptability for the highest ranks. Benslimane receives 72 % acceptability for the first rank, 23 % for the second rank, 4 % for the third rank, 1 % for the fourth rank, and 0 for the ranks 5–9. This means that Benslimane is a robust choice subject to many different possible preferences. Also Casablanca with 25 % acceptability for the first rank and 33 % for the second rank is a possible choice subject to suitable preferences. However, Casablanca is not as robust subject to different preferences, because it can obtain also all other ranks with some probability.

The rank acceptability indices can also be used for ***eliminating some of the worst alternatives***. Among the less acceptable alternatives, in particular Oujda receives either the last or next to last rank with 90 % probability. Eliminating Oujda from the set of best alternatives would a robust choice.

The acceptability profile will provide only a ***rough ranking*** of the alternatives because there is no objective way to combine acceptability indices for different ranks to reach a complete ranking. For forming a complete ranking, Lahdelma and Salminen [8] suggested the *holistic acceptability index*, which is a weighted sum of the rank acceptability indices for different ranks. However, the holistic acceptability index depends on meta-weights in the weighted sum, and meta-weights are subjective. Another problem with using the acceptability indices to form a complete ranking is that if alternatives are removed from or added to the problem, acceptability indices may change, and the mutual order of alternatives may change. This is known as the *rank reversal problem*, present in several MCDA methods. In SMAA the above ranking problems can be resolved by the pairwise winning index, which is presented next.

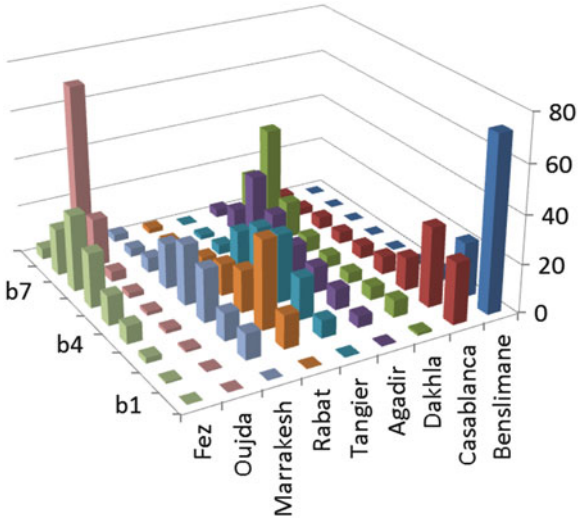


Fig. 1.5: Acceptability profile for alternatives in air cargo hub case

1.3.2 Pairwise Winning Indices

The *pairwise winning index* c_{ik} is the probability for alternative x_i being more preferred than x_k, considering the uncertainty in criteria and preferences [20]. The pairwise winning index is estimated from the simulation statistics as

$$c_{ik} \approx C_{ik}/K \tag{1.4}$$

The pairwise winning indices are useful when comparing the mutual performance of two alternatives. This information can be used e.g. when it is necessary to eliminate inferior alternatives that are dominated by other alternatives.

Unlike the rank acceptability index, the pairwise winning index between one pair of alternatives is independent on the other alternatives. This means that the pairwise winning index can be used to form a ranking among the alternatives. The ranking is obtained by ordering the alternatives so that each alternative x_i precedes all alternatives x_k for which $c_{ik} > 50\,\%$ or some bigger threshold value.

Table 1.4 shows the pairwise winning indices for the air cargo hub case. In this table the alternatives have been ordered to form a complete ranking, which means that all pairwise winning indices in the upper triangle are $>50\,\%$ and $<50\,\%$ in the lower triangle. Observe that there are problems where a complete ranking cannot be obtained. For example, three or more alternatives may win each other in a cyclic manner. In that case such subsets of alternatives obtain the same rank.

Table 1.4: Pairwise winning indices for air cargo hub case (complete ranking)

Alt	Benslimane	Casablanca	Rabat	Tangier	Marrakesh	Agadir	Fez	Dakhla	Oujda
Benslimane	-	74	98	97	100	98	100	99	100
Casablanca	26	-	72	78	77	82	91	81	94
Rabat	2	28	-	62	69	72	90	80	94
Tangier	3	22	38	-	54	78	82	81	96
Marrakesh	0.2	23	31	46	-	69	80	82	93
Agadir	2	18	28	22	31	-	58	72	93
Fez	0.01	9	10	18	20	42	-	64	91
Dakhla	1	19	20	19	18	28	36	-	83
Oujda	0.1	6	6	4	7	7	9	17	-

1.3.3 Central Weight Vectors

The *central weight vector* $\mathbf{w}_i^c$ is the expected center of gravity of the weights that make an alternative most preferred. The central weight vector represents the preferences of a 'typical' DM supporting an alternative. The central weight vectors can be presented to the DMs in order to help them understand how different weights correspond to different alternative choices. To justify their decision, the DMs can, instead of expressing their own trade-off weights for the different criteria, judge if they are willing to accept the central weights of some alternative. The central weight vector for an alternative is estimated from the simulation statistics as

$$\mathbf{w}_i^c \approx \mathbf{W}_i / B_{i1} \tag{1.5}$$

Figure 1.6 (and Table 1.5) shows the central weight vectors for the air cargo hub case. The central weight vector for Fez is not defined, because Fez is an inefficient alternative (first rank acceptability index is zero). For the remaining alternatives the central weight vectors reveal what kind of preferences favor each alternative. For example, Benslimane, which is the most widely acceptable alternative, is most preferred with relatively uniform weights for each criterion. In contrast, Oujda, which is a nearly inefficient alternative, would require very much weight (68 %) on the INVEST criterion alone, and very little weight (2 %) on the POTENTIAL and ENVIRON criteria.

1.3.4 Confidence Factors

The *confidence factor* p_i^c is the probability for an alternative to obtain the first rank when its central weight vector is chosen. The confidence factors measure how robust choice for the first rank an alternative can be if the DMs accept the central weight vector to represent their preferences. A second simulation, presented in Algorithm 2 below, is needed to compute the confidence factors from collected statistics: P_i. The number of times alternative xi was most preferred using weights $\mathbf{w}_i^c$.

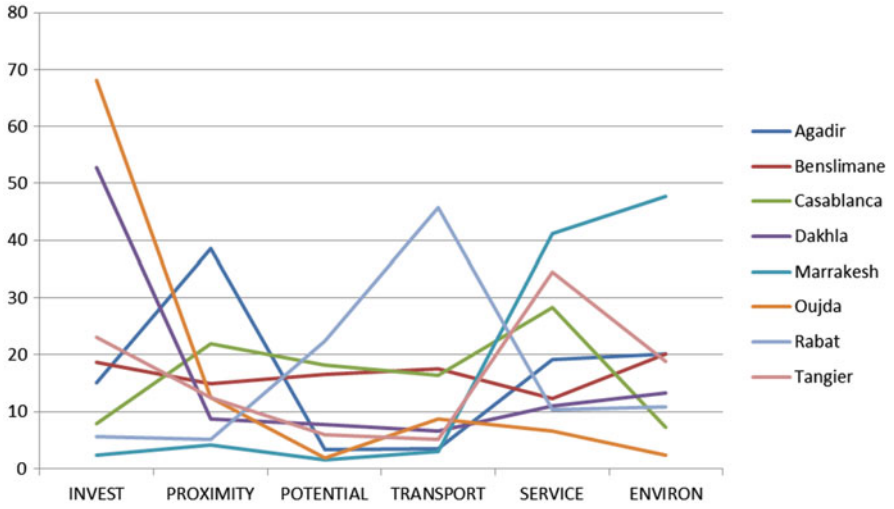


Fig. 1.6: Central weight vectors for air cargo hub case

Algorithm 2. Computation of confidence factors in SMAA

```
Repeat K times {
        Draw x from its distribution
        For the central weight vector w_i^c of each alternative {
                Rank the alternatives using u(x_i, w_i^c)
                Update statistics (P_i) about alternatives
        }
}
```

The confidence factor is estimated from the simulation results as

$$p_i^c \approx P_i/K \tag{1.6}$$

If the confidence factors for all alternatives are low, it means that the criteria measurements are not accurate enough for discriminating the alternatives robustly. In such a situation, collecting more accurate preference information is not sufficient; instead the criteria should be measured more accurately. In the opposite case, when some alternatives have high confidence factors, but low acceptability indices for the best ranks, collecting more accurate preference information may be sufficient.

Table 1.5 presents the confidence factors and corresponding central weight vectors for the alternatives in the air cargo hub case. We can see that only Benslimane and Casablanca are robust choices with suitable preferences falling at or near their central weight vectors. The remaining alternatives are very unlikely to be most preferred even with their central weight vectors. Choosing any of the remaining alternatives would require, besides favorable weights, also more accurate criteria measurement and a new analysis to reassess their robustness.

Table 1.5: Confidence factors and central weights for alternatives in air cargo hub case

Alt	pc	INVEST	PROXIMITY	POTENTIAL	TRANSPORT	SERVICE	ENVIRON
Agadir	6.80	15	39	3	4	19	20
Benslimane	99.98	19	15	17	17	12	20
Casablanca	96.87	8	22	18	16	28	7
Dakhla	17.10	53	9	8	7	11	13
Fez	–	–	–	–	–	–	–
Marrakesh	21.68	2	4	2	3	41	48
Oujda	3.81	68	13	2	9	7	2
Rabat	1.68	6	5	22	46	10	11
Tangier	3.38	23	13	6	5	35	19

1.4 Robustness with Respect to Model Structure

SMAA can be used to analyze the robustness of the decision problem with respect the structure of the decision model. For example, robustness with respect to linearity assumptions in utility/value functions can be analyzed by choosing a more general parametrized utility function and exploring how the solutions change as a function of the degree of non-linearity [14]. As an example, we consider additive utility functions (1.1) where the partial utility functions $u_j(\cdot)$ are non-linear, exponential functions (similar to the Constant Absolute Risk Aversion (CARA) model):

$$u_j(x_j) = \frac{1 - e^{-cx_j}}{1 - e^{-c}} \tag{1.7}$$

The parameter c measures the curvature of the function. Positive values of c result in concave shapes and negative values yield convex shapes. When $c \to 0$, the function approaches a linear function.

Partial utility functions with positive curvature compose into an overall utility function favoring alternatives that are uniformly good on each criterion. Negative curvature favors alternatives that are superior on any single criterion. In any case, a dominated alternative can never be the most preferred.

To analyze the robustness of the air cargo hub case, we study how the first rank acceptability indices (b_i^1) and lexicographic ranks of alternatives depend on the curvature of the partial utility functions. For the cardinally measured criteria (INVEST, POTENTIAL, TRANSPORT, SERVICE) we consider 11 curvature levels: $c \in \{-8, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 8\}$. Figure 1.7 illustrates the corresponding partial utility functions. The curvature for $c = 8$ is very high; the marginal value at $x_j = 0$ is 2980 times higher than at 1. The different partial utility functions may have different shapes. In this example we consider only the situation where each cardinal criterion has the same curvature.

In the following we analyze how much the acceptability indices and the lexicographic rankings of alternatives change when moving from the linear model to

each of the non-linear models. Table 1.6 shows that the acceptability indices are very robust subject to small non-linearities. Significant (>5 %) changes in acceptability indices occur only for Benslimane and Casablanca at $c > 2$, for Benslimane at $c < -1$, for Casablanca at $c < -2$, and for Dakhla at $c < -4$.

Table 1.7 shows that the lexicographic ranking of the top alternatives is very robust subject to non-linearity. Benslimane and Casablanca preserve their first and second rank regardless the curvature. Dakhla preserves its third rank for negative curvature but for positive curvature it loses its position.

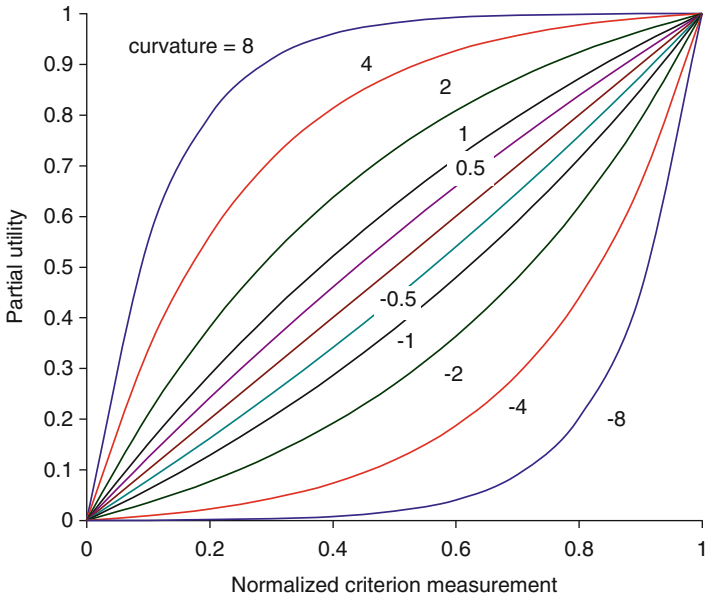


Fig. 1.7: Partial utility functions with different amounts of non-linearity

1.5 Recent Developments of SMAA

Recent developments of SMAA include more efficient computational methods and extensions to different decision models.

In most cases the SMAA computations can be performed very efficiently using straight forward Monte Carlo simulation. However, the computation may slow down in case of complex preference information. In such cases, the Markov Chain Monte Carlo (MCMC) simulation technique can be used to speed up the computation [27]. The JSMAA open source implementation of SMAA includes the MCMC technique and performs the simulation as a background process while the user views or edits the model (see www.smaa.fi, [23]).

Table 1.6: Acceptability indices (%) of alternatives with different amount of curvature. Over 5 % changes highlighted for illustrative purposes

	Curvature c										
Alternative	−8	−4	−2	−1	−0.5	0	0.5	1	2	4	8
Agadir	0.05	0.09	0.16	0.26	0.31	0.40	0.48	0.53	0.75	1.40	2.00
Benslimane	**47.00**	**54.00**	**62.00**	**67.00**	70.00	72.00	75.00	76.00	**79.00**	**82.00**	**85.00**
Casablanca	25.00	30.00	**31.00**	29.00	28.00	25.00	24.00	22.00	**18.00**	**12.00**	**9.00**
Dakhla	**23.00**	**13.00**	5.80	2.90	1.70	1.00	0.53	0.17	0.01	0.00	0.00
Fez	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Marrakesh	0.00	0.00	0.01	0.01	0.01	0.03	0.04	0.07	0.14	0.72	1.00
Oujda	0.84	0.41	0.19	0.09	0.05	0.02	0.01	0.01	0.00	0.00	0.00
Rabat	3.60	2.00	0.99	0.52	0.43	0.28	0.15	0.14	0.16	0.19	0.08
Tangier	0.00	0.00	0.05	0.12	0.22	0.38	0.54	0.73	1.50	2.80	2.80

Table 1.7: Lexicographic ranks of alternatives with different amount of curvature

	Curvature c										
Alternative	−8	−4	−2	−1	−0.5	0	0.5	1	2	4	8
Agadir	6	6	6	5	5	4	5	4	4	4	4
Benslimane	1	1	1	1	1	1	1	1	1	1	1
Casablanca	2	2	2	2	2	2	2	2	2	2	2
Dakhla	3	3	3	3	3	3	4	5	7	8	8
Fez	9	9	9	9	9	9	9	9	8	7	7
Marrakesh	7	7	8	8	8	7	7	7	6	5	5
Oujda	5	5	5	7	7	8	8	8	9	9	9
Rabat	4	4	4	4	4	6	6	6	5	6	6
Tangier	8	8	7	6	6	5	3	3	3	3	3

Extensions to different decision models include different shaped utility or value functions and also decision models not based on utility functions. Cohen et al. [3] applied SMAA with an additive value function where the partial value functions (marginal value functions) were piecewise linear monotonic mappings. They varied the mappings during simulation using a random process resembling treatment of ordinal criteria measurements in SMAA. Babalos et al. [2] applied the SMAA-2 framework and considered three different aggregate evaluation measures: the holistic acceptability index, Borda count method, and average score. Kontu et al. [7] extended the SMAA method to handle a hierarchy of criteria and sub-criteria. A criteria hierarchy is useful when the number of criteria is large.

Additive utility function models assume independence between criteria. SMAA with Choquet integral by Angilella et al. [1] considers interaction between criteria. The Choquet integral can be seen as a value function where positive or negative interaction between criteria is also contributing to the evaluation of alternatives. The Choquet integral is thus a more general decision model than the additive value function. Lahdelma and Salminen [14] studied the robustness of decision problems with respect to the shape of the utility function, as demonstrated in the previous section.

The SMAA-PROMETHEE method by Corrente et al. [4] is a recent extension of SMAA to non-utility function based methods. PROMETHEE is based on an outranking procedure where fuzzy preference relations between alternatives are aggregated together to yield a partial order (PROMETHEE I) or complete order (PROMETHEE II). Durbach et al. [6] extended the analytic hierarchy process (AHP) to consider imprecise or uncertain pairwise comparisons by probability distributions. The resulting SMAA-AHP method is suitable for group decision making problems, where it is difficult to agree on precise pairwise comparisons. Wang et al. [28] extended the Complementary Judgement Matrix (CJM) method in a similar manner. CJM differs from AHP in the way how the pairwise comparisons are expressed, and in how the weights are solved from inconsistent comparisons. In particular, the weights in CJM are determined by minimizing the square sum of inconsistency errors.

1.6 Discussion

In SMAA uniform distributions are used to represent absence of information both in criteria and preferences. Ordinal criteria are transformed into cardinal measurements by simulating consistent ordinal to cardinal mappings. The simulation process is equivalent to treating the absence of interval information of ordinal scales as uniform joint distributions. Similarly, absence of weight information is treated as a uniform joint distribution in the feasible weight space.

Although SMAA can be used with arbitrarily shaped utility functions, in real-life applications simple forms, such as linear or some concave shapes are most commonly applied. Assessing the precise preference structure of DMs can be difficult and time-consuming in practice. SMAA can be used to test the robustness of the problem also with respect to the decision model, as illustrated in the previous section. If the problem can be identified as robust with respect to model structure, it may be possible to assume a simpler model in the interaction between the DMs.

The strength of SMAA in robustness analysis of multicriteria decision aiding problems is that it is able to handle the whole range of uncertain, imprecise or partially missing information flexibly using suitable probability distributions. Typically, a real-life decision process may start with very vague and uncertain criteria and preference information. The information will become gradually more accurate during the process. SMAA can be used in such processes repeatedly after any refinement of information, until a robust decision can be identified and agreed on. SMAA reveals if the information is accurate enough for making the decision, and also pinpoints which parts of the information need to be refined. This can (1) protect the DMs from making wrong decisions based on insufficient information and also (2) cause significant savings in information collection if less accurate information is sufficient for making a robust decision.

References

1. Angilella, S., Corrente, S., Greco, S.: Stochastic multiobjective acceptability analysis for the Choquet integral preference model and the scale construction problem. Eur. J. Oper. Res. **240**(1), 172–182 (2015)
2. Babalos, V., Philippas, N., Doumpos, M., Zopounidis, C.: Mutual funds performance appraisal using stochastic multicriteria acceptability analysis. Appl. Math. Comput. **218**(9), 5693–5703 (2012)
3. Cohen, S., Doumpos, M., Neofytou, E., Zopounidis, C.: Assessing financial distress where bankruptcy is not an option: an alternative approach for local municipalities. Eur. J. Oper. Res. **218**(1), 270–279 (2012)
4. Corrente, S., Figueira, J.R., Greco, S.: The SMAA-PROMETHEE method. Eur. J. Oper. Res. **239**(2), 514–522 (2014)
5. David, H.A., Nagaraja, H.N.: Order statistics. Wiley Online Library (1970)
6. Durbach, I., Lahdelma, R., Salminen, P.: The analytic hierarchy process with stochastic judgements. Eur. J. Oper. Res. **238**(2), 552–559 (2014)
7. Kontu, K., Rinne, S., Olkkonen, V., Lahdelma, R., Salminen, P.: Multicriteria evaluation of heating choices for a new sustainable residential area. Energy Build. **93**, 169–179 (2015)
8. Lahdelma, R., Salminen, P.: SMAA-2: stochastic multicriteria acceptability analysis for group decision making. Oper. Res. **49**(3), 444–454 (2001)
9. Lahdelma, R., Salminen, P.: Pseudo-criteria versus linear utility function in stochastic multicriteria acceptability analysis. Eur. J. Oper. Res. **141**(2), 454–469 (2002)
10. Lahdelma, R., Salminen, P.: Stochastic multicriteria acceptability analysis using the data envelopment model. Eur. J. Oper. Res. **170**(1), 241–252 (2006)
11. Lahdelma, R., Salminen, P.: Prospect theory and stochastic multicriteria acceptability analysis (SMAA). Omega **37**(5), 961–971 (2009)
12. Lahdelma, R., Salminen, P.: A method for ordinal classification in multicriteria decision making. In: Proceedings of the 10th IASTED International Conference on Artificial Intelligence and Applications, vol. 674, pp. 420–425 (2010)
13. Lahdelma, R., Salminen, P.: Stochastic multicriteria acceptability analysis (SMAA). In: Ehrgott, M., Figueira, J., Greco, S. (eds.) Trends in Multiple Criteria Decision Analysis. International Series in Operations Research and Management Science, vol. 142, pp. 285–316. Springer, Berlin (2010)
14. Lahdelma, R., Salminen, P.: The shape of the utility or value function in stochastic multicriteria acceptability analysis. OR Spectr. **34**(4), 785–802 (2012)
15. Lahdelma, R., Hokkanen, J., Salminen, P.: SMAA-stochastic multiobjective acceptability analysis. Eur. J. Oper. Res. **106**(1), 137–143 (1998)
16. Lahdelma, R., Miettinen, K., Salminen, P.: Ordinal criteria in stochastic multicriteria acceptability analysis (SMAA). Eur. J. Oper. Res. **147**(1), 117–127 (2003)
17. Lahdelma, R., Miettinen, K., Salminen, P.: Reference point approach for multiple decision makers. Eur. J. Oper. Res. **164**(3), 785–791 (2005)
18. Lahdelma, R., Makkonen, S., Salminen, P.: Multivariate Gaussian criteria in SMAA. Eur. J. Oper. Res. **170**(3), 957–970 (2006)
19. Lahdelma, R., Makkonen, S., Salminen, P.: Two ways to handle dependent uncertainties in multi-criteria decision problems. Omega **37**(1), 79–92 (2009)
20. Leskinen, P., Viitanen, J., Kangas, A., Kangas, J.: Alternatives to incorporate uncertainty and risk attitude in multicriteria evaluation of forest plans. For. Sci. **52**(3), 304–312 (2006)
21. Menou, A., Benallou, A., Lahdelma, R., Salminen, P.: Decision support for centralizing cargo at a Moroccan airport hub using stochastic multicriteria acceptability analysis. Eur. J. Oper. Res. **204**(3), 621–629 (2010)
22. Taylor, B.N.: Guidelines for evaluating and expressing the uncertainty of NIST measurement results. Tech. Rep. 1297, National Institute of Standards and Technology (1994)
23. Tervonen, T.: JSMAA: open source software for SMAA computations. Int. J. Syst. Sci. **45**(1), 69–81 (2014)

24. Tervonen, T., Figueira, J.R.: A survey on stochastic multicriteria acceptability analysis methods. J. Multi-Criteria Decis. Anal. **15**(1–2), 1–14 (2008)
25. Tervonen, T., Lahdelma, R.: Implementing stochastic multicriteria acceptability analysis. Eur. J. Oper. Res. **178**(2), 500–513 (2007)
26. Tervonen, T., Figueira, J.R., Lahdelma, R., Dias, J.A., Salminen, P.: A stochastic method for robustness analysis in sorting problems. Eur. J. Oper. Res. **192**(1), 236–242 (2009)
27. Tervonen, T., van Valkenhoef, G., Baştürk, N., Postmus, D.: Hit-and-run enables efficient weight generation for simulation-based multiple criteria decision analysis. Eur. J. Oper. Res. **224**(3), 552–559 (2013)
28. Wang, H., Jiao, W., Lahdelma, R., Zhu, C., Zou, P.: Stochastic multicriteria acceptability analysis for evaluation of combined heat and power units. Energies **8**(1), 59–78 (2015)
29. Yevseyeva, I.: Solving classification problems with multicriteria decision aiding approaches. Ph.D. thesis, Jyväskylä Studies in Computing 84 (2007)

Chapter 2
Data-Driven Robustness Analysis for Multicriteria Classification Problems Using Preference Disaggregation Approaches

Michael Doumpos and Constantin Zopounidis

Abstract The preference disaggregation framework of multicriteria decision aid focuses on inferring decision models from data. In this context, the robustness of the results is of major importance to ensure that quality recommendations are provided. In this chapter we examine this issue adopting a data-driven perspective, focusing on the effect due to changes in the data used for model construction. The analysis is implemented for decision models expressed in the form of additive value functions for multicriteria classification problems. Simple analytic measures are introduced based on well-known optimization tools. The proposed measures enrich existing robust multicriteria approaches with additional information taken directly from the available data though an analytical approach. The properties and performance of the new robustness indicators are illustrated through their application to an example data set.

2.1 Introduction

Multiple criteria decision aid (MCDA) is involved with supporting the structuring and modeling of decision problems involving multiple conflicting criteria. Similarly to other operations research/management science approaches, MCDA methods are

M. Doumpos (✉)
School of Production Engineering and Management, Financial Engineering Laboratory, Technical University of Crete, University Campus, 73100 Chania, Greece
e-mail: mdoumpos@dpem.tuc.gr

C. Zopounidis
School of Production Engineering and Management, Financial Engineering Laboratory, Technical University of Crete, University Campus, 73100 Chania, Greece

Audencia Business School, Nantes, France
e-mail: kostas@dpem.tuc.gr

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_2

also based on modeling assumptions, related to the characteristics of the problem, the aggregation of the decision criteria, and the preferential system of the decision maker (DM). Naturally, these assumptions incorporate uncertainties, fuzziness, and errors, thus affecting the quality of the obtained recommendations. Thus, changes in the decision context and the available data may lead to completely different outputs.

In this framework, robustness analysis has emerged as a major research issue in MCDA, emphasizing the need to re-think the traditional multicriteria framework aiming towards providing satisfactory recommendations even in cases where the decision context is altered. Roy [21] described in detail the *robustness concern*, arguing that it is raised by *vague approximations* and *zones of ignorance* that cause the formal representation of a problem to diverge from the real-life context, due to: (1) the way imperfect knowledge is treated, (2) the inappropriate preferential interpretation of certain types of data (e.g., transformations of qualitative attributes), (3) the use of modeling parameters to grasp complex aspects of reality, and (4) the introduction of technical parameters with no concrete meaning.

MCDA provides a wide arsenal of methodologies and techniques that enable the systematic treatment of decision problems under multiple criteria. In this chapter we focus on the preference disaggregation approach (PDA), which is involved with the inference of preferential information and decision models from data [15]. PDA techniques can greatly facilitate the model construction process, reducing the cognitive effort required by DMs when specifying complex preferential information and modeling parameters.

Robustness analysis in the framework of PDA is based on analytic and simulation techniques (for an overview see [7]). This chapter considers the former approach, which is based on two main schemes. The first focuses on the construction of a single decision model that best represents the available decision instances [5, 13], whereas the second is involved with the formulation of a range of recommendations on the basis of all models compatible with the given data [10, 12]. In this chapter we re-analyze the robustness of such approaches and introduce new robustness metrics following a data-driven perspective. More specifically, we are concerned with robustness issues in terms of variations in the data instances used to infer a decision model. A similar view of robustness is very common on other fields also involved with model inference from data (e.g., statistical learning [6]), but its analytic treatment in the context of MCDA has been limited so far, despite the existence of experimental results supporting its significance [8, 24]. This chapter contributes in that direction and proposes tools based on well-known concepts from optimization theory. The analysis is focused on decision models expressed in the form of additive value functions for classification (sorting) problems, which involve the assignment of a finite set of alternative options into predefined performance categories [27]. For the purposes of the presentation an illustrative example is used.

The rest of the chapter is organized as four sections. Section 2.2 introduces the framework of preference disaggregation analysis for classification problems and presents the main existing robustness analysis techniques and approaches from the MCDA literature. Section 2.3 discusses the importance of the proposed data-driven framework for robustness analysis in disaggregation techniques and introduces

new robustness indicators constructed on the basis of this framework. Section 2.4 presents results from the application on an example data set and finally Sect. 2.5 concludes the chapter and discusses some future research directions.

2.2 Preference Disaggregation for Multicriteria Classification

2.2.1 General Framework

Multicriteria problems involve multi-objective optimization and discrete evaluation cases. In this chapter we are concerned with the latter type, which is about the evaluation of a set X of discrete alternatives over n performance criteria. The result of the evaluation may be expressed in different forms, such as a choice, ranking, and classification. The present study focuses on classification problems, where the alternatives under consideration should be classified into q rank-ordered performance categories $C_1 \succ C_2 \succ \cdots \succ C_q$. Category C_1 is assumed to consist of the best alternatives whereas C_q consists of the worst performing ones.

In this context, a decision model $F(\mathbf{x}, \beta) \rightarrow \{C_1, \ldots, C_q\}$ aggregates the available information about the criteria and provides recommendations about the classification of the alternatives. The model is explicitly defined by the parameters β, which may relate to the relative importance of the criteria or other information about the aggregation process.

In the field of MCDA there is a wide range of different types of decision and evaluation models. Some common examples include value functions [17], outranking models [20, 25], and decision rules [9]. Bouyssou et al. [2] provide a comprehensive overview of different MCDA models and their characterization.

For the reminder of the presentation this chapter will focus on additive value function (AVF) models, which have been widely used in MCDA. The general form of an AVF is:

$$V(\mathbf{x}_i) = \sum_{k=1}^{n} w_k v_k(x_{ik}) \tag{2.1}$$

where $\mathbf{x}_i = (x_{i1}, \ldots, x_{in})$ is the data vector for alternative i (x_{ij} being the data of i on criterion j), $w_1, \ldots, w_n \geq 0$ are trade-off constants (normalized to sum up to one) representing the relative importance of the criteria, and $v_1(\cdot), \ldots, v_n(\cdot)$ are the marginal value functions of the criteria. The marginal value functions decompose the overall performance $V(\mathbf{x}_i)$ of each alternative i into partial assessments at the criteria level, each usually scaled between 0 and 1.

The most straightforward approach to use a value function model to classify an alternative into predefined rank-ordered classes, is to employ the following decision rule:

$$t_\ell < V(\mathbf{x}_i) < t_{\ell-1} \Leftrightarrow \mathbf{x}_i \in C_\ell \tag{2.2}$$

where $t_0 = 1 > t_1 > t_2 > \cdots > t_{q-1} > t_q = 0$ are thresholds that distinguish the classes. Alternative classification rules can also be employed such as the example-based approach of Greco et al. [12] or the hierarchical model of Zopounidis and Doumpos [26].

In the framework of PDA, the parameters of the model are inferred from a sample of m decision instances $X' = \{\mathbf{x}_i, y_i\}_{i=1}^{m}$, where y_i denotes the given class label for alternative i. This sample (referred to as the reference set) may consist of decisions about alternatives considered in past situations or decisions about a set of alternatives which can be easily judged by the DM [15].

Formally, the model that is most compatible with the information in the reference set is defined by parameters $\widehat{\beta}^*$ such that:

$$\widehat{\beta}^* = \arg\min_{\widehat{\beta} \in \mathscr{A}} L[Y_{X'}, F(X', \widehat{\beta})] \tag{2.3}$$

where $F(X', \widehat{\beta})$ denotes the outputs of a model with parameters $\widehat{\beta}$ for the alternatives in X', $\mathscr{A}$ is the set of acceptable parameter values, and $L(\cdot)$ is a function that measures the differences between the recommendations of the model and the actual assessments $Y_{X'}$ for the reference alternatives. If the solution of the above problem (2.3) is judged satisfactory, then the inferred parameters $\widehat{\beta}^*$ can be used to extrapolate the model to any other alternative outside the reference set.

For a value function model, problem (2.3) is expressed in a mathematical programming form. In particular, the inference of a classification model (weights of the criteria, marginal value functions, and classification thresholds) from the reference examples can be expressed as the following optimization problem:

$$\text{min} \quad \sum_{\ell=1}^{q} \frac{1}{m_\ell} \sum_{\mathbf{x}_i \in C_\ell} (\sigma_i^+ + \sigma_i^-) \tag{2.4}$$

$$\text{s.t.} \quad V(\mathbf{x}_i) + \sigma_i^+ \geq t_\ell + \delta \qquad \forall \mathbf{x}_i \in C_\ell, \ell = 1, \ldots, q-1 \tag{2.5}$$

$$V(\mathbf{x}_i) - \sigma_i^- \leq t_\ell - \delta \qquad \forall \mathbf{x}_i \in C_\ell, \ell = 2, \ldots, q \tag{2.6}$$

$$t_\ell - t_{\ell+1} \geq \varepsilon \qquad \ell = 1, \ldots, q-2 \tag{2.7}$$

$$V(\mathbf{x}_*) = 0,\ V(\mathbf{x}^*) = 1 \tag{2.8}$$

$$V(\mathbf{x}) \geq V(\mathbf{x}') \qquad \forall \mathbf{x} \geq \mathbf{x}' \tag{2.9}$$

$$\sigma_i^+, \sigma_i^- \geq 0 \qquad i = 1, \ldots, m \tag{2.10}$$

The objective function minimizes the total weighted classification error, where the weights are defined on the basis of the number of reference alternatives from each class $(m_1, \ldots, m_q)$. The error variables σ^+ and σ^- are defined through constraints (2.5)–(2.6) as the magnitude of the violations of the classification rules (2.2) (δ is a small positive constant used to ensure the string inequalities), whereas constraint (2.7) ensures that the class thresholds are defined in a decreasing sequence (ε is a small positive constant). Constraint (2.8) defines the scale of the additive model between 0 and 1 (0 corresponds to the performance of the least preferred

alternative $\mathbf{x}_*$ and 1 corresponds to the performance of an ideal action $\mathbf{x}^*$). Finally, constraint (2.9) ensures that the model is non-decreasing with respect to the performance criteria (assuming all criteria are in maximization form).

For the case of an AVF, the above optimization problems can be written in linear programming form with a piece-wise linear modeling of the marginal values function (for the modeling details, see [4, 14]).

2.2.2 Robust Approaches

The robustness concern in the context of PDA arises because often alternative decision models can be inferred in accordance with the information embodied in the set of reference decision examples that a DM provides (i.e., the optimization model (2.4)–(2.10) often has multiple optimal solutions). This is particularly true for reference sets that do not contain inconsistencies, but it is also relevant when inconsistencies exist (in the PDA context, inconsistencies are usually resolved algorithmically or interactively with the DM before the final model is built; see for instance [19]).

With a consistent reference set the error variables can be removed from formulation (2.4)–(2.10), which then reduces to a set of feasible linear constraints defining all acceptable models that are compatible with the assignment of the reference alternatives.

$$\begin{aligned}
&V(\mathbf{x}_i) \geq t_\ell + \delta && \forall \mathbf{x}_i \in G_\ell,\ \ell = 1,\ldots,q-1 \\
&V(\mathbf{x}_i) \leq t_\ell - \delta && \forall \mathbf{x}_i \in G_\ell,\ \ell = 2,\ldots,q \\
&t_\ell - t_{\ell+1} \geq \varepsilon && \ell = 1,\ldots,q-2 \\
&V(\mathbf{x}_*) = 0,\ V(\mathbf{x}^*) = 1 && \\
&V(\mathbf{x}) \geq V(\mathbf{x}') && \forall \mathbf{x} \geq \mathbf{x}'
\end{aligned} \tag{2.11}$$

The size of the polyhedron defined through (2.11) is associated with the robustness of the results and can be affected by a number of factors. The most important of these factors relate to the adequacy of the set of reference examples and the complexity of the selected decision modeling form. The former is immediately related to the quality of the information on which model inference is based. Vetschera et al. [24] performed an experimental analysis to investigate how the size of the reference set affects the robustness and accuracy of the resulting multicriteria models in classification problems. They found that small reference sets (e.g., with a limited number of alternatives with respect to the number of criteria) lead to decision models that are neither robustness nor accurate. Expect for its size, other characteristics of the reference set are also relevant, such as the existence of noisy data, outliers, the existence of correlated criteria, etc. [4].

Traditional disaggregation techniques such as the family of the UTA methods use linear programming post-optimality techniques [22] in order to build a representative AVF defined as the average solution of some characteristic extreme points of the feasible polyhedron (2.11). Other approaches for selecting the most representative

decision model include the regularization approach of Doumpos and Zopounidis [5], the analytic center formulation of Bous et al. [1], and the max-min model of Greco et al. [13]. As explained by Doumpos et al. [8] such approaches seek to identify (analytically) central solutions to the polyhedron defined by (2.11), which are expected to be more robust to changes in the data and the setting of the analysis.

Recently, alternative approaches have been proposed that enable the formulation of recommendations based on multiple decision models. Two main schemes can be identified in this framework. The first is based on simulation techniques, which are based on sampling, at random, different solutions (value functions) from the polyhedron defined by (2.11). The simulation process provides an approximate description of all models compatible with the classifications for the reference set and enables the formulation of a range of recommendations associated with probabilistic measures of confidence (see, for instance, [23]).

The second scheme, on which this study is focused, is based on approaches that seek to characterize the full set of acceptable models through analytic techniques, rather than using simulation. In particular, Greco et al. [12] introduced a modeling framework that takes into account all decision models (AVFs) compatible with the constraints (2.11). Their approach is based on the definition of necessary and possible assignments. The set of necessary assignments $\mathscr{N}_j$ for a non-reference alternative $j \notin X'$ consists of the classes in which j is classified by all models compatible with the reference set, whereas the set of possible assignments $\mathscr{P}_j$ includes the results supported by at least one decision model. Obviously, $\mathscr{N}_j \subseteq \mathscr{P}_j$. Furthermore, it should be noted that these definitions cover the general case where the reference alternatives might be classified in multiple classes (rather than the specific case described above where each alternative is assigned into only one class, in which case $\mathscr{N}_j$ is either empty or singleton).

Figure 2.1 provides a graphical illustration of the necessary and possible assignments for a two-class problem, assuming a linear decision model (linear value function). With the given reference set consisting of alternatives classified in two categories (circles and rectangles), it is evident that all models that separate the two classes assign the non-reference alternative $\mathbf{x}_1$ into class C_1. On the other hand, the precise classification of the non-reference action $\mathbf{x}_2$ is not possible. In fact, this alternative can be assigned to any of the two categories.

The necessary and possible assignments for a non-reference alternative j can be obtained through linear programming [12, 16]. In particular, a class C_ℓ belongs to the set of possible assignments for a non-reference alternative j if the optimal objective value of the following linear program is strictly positive:

$$\begin{array}{ll} \max & \gamma \\ \text{s.t.} & t_\ell + \gamma \leq V(\mathbf{x}_j) \leq t_{\ell+1} - \gamma \\ & \text{constraints (2.11) for } X' \end{array} \tag{2.12}$$

Similarly, a class C_ℓ belongs to the set of necessary assignments for alternative j if either of the following two linear programs has a non-positive optimal objective function value:

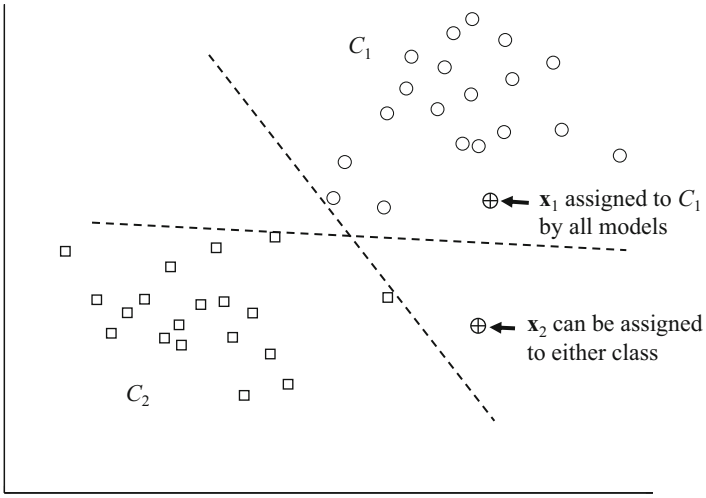


Fig. 2.1: An illustration of possible and necessary assignments

$$\begin{array}{lll} \max & \gamma & \\ \text{s.t.} & V(\mathbf{x}_j) \geq t_{\ell-1} + \gamma & \\ & \text{constraints (2.11) for } X' & \end{array} \qquad \begin{array}{lll} \max & \gamma & \\ \text{s.t.} & V(\mathbf{x}_j) \leq t_{\ell} - \gamma & \\ & \text{constraints (2.11) for } X' & \end{array} \tag{2.13}$$

If $\gamma \leq 0$ in the optimal solution of the left problem, then j cannot be assigned to any of the classes in the set $\{C_1, \ldots, C_{\ell-1}\}$, which implies that $C_\ell \in \mathscr{N}_j$. On the other hand, if the optimal solution of the right problem yields $\gamma \leq 0$, then j cannot be assigned to any of the classes in the set $\{C_{\ell+1}, \ldots, C_q\}$, which again implies that $C_\ell \in \mathscr{N}_j$.

It follows that, for every non-reference alternative j, the obtained possible assignments define a range $[L_j, U_j]$ with the worst and best possible ratings that can be defined on the basis of the information available in the evaluations of the reference actions.

The identification of the necessary and possible assignments provides valuable additional information as opposed to simple point recommendations obtained from a single decision model, thus enhancing the robustness of the results. However, given that the necessary and possible assignments are data-driven results (i.e., they are obtained from a specific reference set), it is apparent that they are also subject to the robustness concern. Figure 2.2 provides an illustration of this issue. According to the given two-class reference set (circles and rectangles), the indicated non-reference alternative is necessarily assigned to class C_2 by all linear value functions compatible with the available reference evaluations. This result, however, is not robust because a reconsideration of the evaluations for the two circled reference alternatives will lead to a different outcome.

Kadziński and Tervonen [16] proposed the combination of robust analytic procedures based on the specification of the necessary and analytic assignments with

simulation techniques. The latter provide further information in probabilistic form about the necessary and possible assignments. Simulation-based methods, however, only provide an approximate description of the problem data and they can be computationally intensive for larger data sets involving many alternatives and criteria.

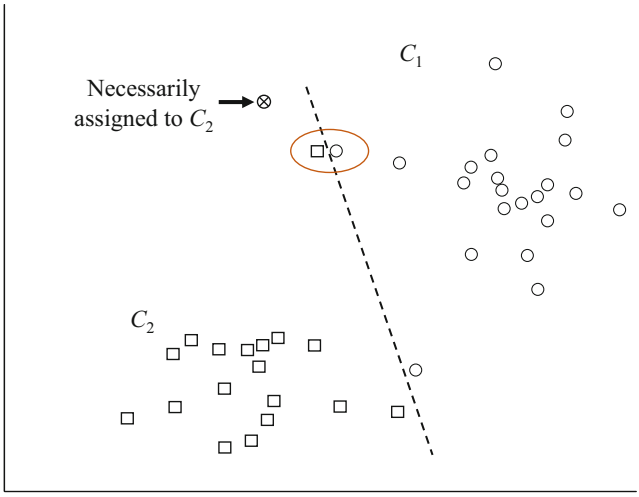


Fig. 2.2: An example of a necessary assignment that is not robust

In the next section we present new ways and metrics to gain further insight into the robustness of necessary and robust assignments, without requiring the use of simulation. The proposed approaches adopt a data-driven perspective, in the sense that they are based on the properties of the available reference set. Their implementation is grounded on well-known techniques from optimization theory.

2.3 Data-Driven Robustness Indicators for Multicriteria Classification Problems

Motivated by the above discussion about the robustness concern for classification recommendations formulated using a set of decision models, this section presents simple techniques that can be used to gain a better understanding of the robustness issue in relation to the problem data, as represented in a set of reference assessments. The main idea is based on the analysis of the changes in the feasible polyhedron (2.11) due to the incorporation of the necessary/possible assignments to a given reference set.

To this end, first a simple support measure can be defined. Assume that according to a given reference set X', a non-reference alternative j can be assigned to any of the classes in the interval $[L_j, U_j]$. Then, the support measure S_j is defined

as the minimum number of changes that need to be made in the assignments of the reference actions in order to allow the classification of j into classes outside $[L_j, U_j]$. The lower this support measure is, the less robust in the obtained interval assignment $[L_j, U_j]$, because minor changes in the reference set will lead to different conclusions.

The computation of support can be done in a straightforward manner through the solution of the following two mixed-integer linear programming problems:

$$\begin{array}{lll} \min & \sum_{i=1}^{m} (\sigma_i^+ + \sigma_i^-) & \\ \text{s.t.} & V(\mathbf{x}_j) \geq t_{L_j-1} + \delta & \\ & \text{constraints (2.5)–(2.9) for } X' & \\ & \sigma_i^+, \sigma_i^- \in \{0,1\} & \end{array} \qquad \begin{array}{lll} \min & \sum_{i=1}^{m} (\sigma_i^+ + \sigma_i^-) & \\ \text{s.t.} & V(\mathbf{x}_j) \leq t_{U_j} - \delta & \\ & \text{constraints (2.5)–(2.9) for } X' & \\ & \sigma_i^+, \sigma_i^- \in \{0,1\} & \end{array} \tag{2.14}$$

The left problem applied to cases where $L_j \geq 2$ and returns the minimum number of changes that need to be made in the assignments of the reference actions in order to classify the non-reference alternative j to the set of categories $\{C_1, \ldots, C_{L_j-1}\}$. Similarly, the right problem applies to cases with $U_j \leq q-1$ and returns the minimum number of changes that need to be made in the assignments of the reference actions in order to classify the non-reference alternative j to the set of categories $\{C_{U_j+1}, \ldots, C_q\}$.

The support measure S_j can then be defined as the minimum of the two objective functions at the optimal solutions of the two problems. When $L_j = 1$ and $U_j = q$, then S_j is by definition equal to zero. In other cases, if S_j is non-zero but low, then the DM may accept the changes identified through the solution of the above optimization models, thus forming a new reference set X'_j.

In order to compare the size of the feasible polyhedron corresponding to the new reference set to the one of the initially available reference set X', we consider two measures based on well-known results from optimization theory.

The first measure is based on the radius of the largest ball inscribed inside the feasible polyhedron. Given a polyhedron $\{\mathbf{x} \,|\, \mathbf{A}^\top \mathbf{x} \leq \mathbf{b}\}$, the radius r of the largest ball inscribed in it can be computed from the following linear program [3]:

$$\begin{array}{ll} \max & r \\ \text{s.t.} & \mathbf{a}_i^\top \mathbf{x} + r\|\mathbf{a}_i\|_2 \leq b_i, \quad \forall i \end{array} \tag{2.15}$$

where $\mathbf{a}_i$ is the ith row of $\mathbf{A}$.

This approach can be straightforwardly applied to find the radius r_0 of the largest ball inscribed inside the polyhedron (2.11) corresponding to the original reference set and compare it to the radius r_j of the largest ball for the modified reference set X'_j. Then, the following robustness measure can be defined:

$$R_j = \frac{\log r_0}{\log r_j} \tag{2.16}$$

The case $R_j > 1$ indicates that the modified reference set X'_j, which allows the classification of the non-reference alternative j outside its first computed range of assignments $[L_j, U_j]$, provides more options for choosing an acceptable decision model. Thus, the modification of X' towards the new reference set X'_j is likely to lead to more robust results. On the other hand, the case $R_j < 1$ indicates that the modified reference set is more restrictive compared to X', which implies that this modification is more sensitive to changes of the reference set (i.e., less robust).

Alternatively to the above metric, the size of the polyhedron corresponding to the set of compatible decision models, can be assessed through the volume of the maximum ellipsoid inscribed inside the polyhedron. Compared to the above metric, this is a more suitable approach for irregular polyhedra, which can not be well described by the largest ball inscribed inside them (e.g., because they have large extremes).

The volume of the largest ellipsoid inside a polyhedron $\{\mathbf{x} \,|\, \mathbf{A}^\top \mathbf{x} \leq \mathbf{b}\}$ can be found from the solution of the following convex optimization problem [3]:

$$\begin{aligned} &\min\ v = \log\det \mathbf{B}^{-1} \\ &\text{s.t.}\quad \|\mathbf{B}\mathbf{a}_i\|_2 + \mathbf{a}_i^\top \mathbf{d} \leq b_i, \qquad \forall i \end{aligned} \tag{2.17}$$

where $\mathbf{d}$ is a vector of decision variables defining the center of the ellipsoid whose volume is proportional to $\det \mathbf{B}$. Similarly to the previous measure, this optimization problem can be used to compare the volume of the largest ellipsoid inscribed inside the polyhedron (2.11) corresponding to the original reference set, against the volume for the modified reference set X'_j. The robustness measure in this case is defined as follows:

$$V_j = \frac{v_0}{v_j} \tag{2.18}$$

Similarly to the interpretation of (2.16), the case $V_j > 1$ indicates that the modification of the original reference set to allow the classification of the non-reference alternative j outside its first computed range of assignments $[L_j, U_j]$, leads to more available options for selecting an acceptable decision model (i.e., higher robustness), versus the case $V_j < 1$, which corresponds to a small (less robust) polyhedron.

2.4 Illustrative Results

In order to examine the potentials of the data-driven robustness measures introduced in the previous section, we present results from their application to a data set taken from Mousseau et al. [18]. The data involve 100 alternatives evaluated on seven criteria (all in minimization form). The alternatives are classified in three performance categories: the high performance class (category H), the medium performance group (category M), and the low performance alternatives (class L).

For the purposes of the analysis, a reference set of 30 randomly selected alternatives (10 alternatives from each category) is used. Table 2.1 presents the results for the necessary ($\mathscr{N}$) and possible ($\mathscr{P}$) assignments of the 70 non-reference

alternatives obtained with the chosen reference set, as opposed to their actual classification (columns). Out of the five alternatives actually belonging in the high performance class, four are assigned to the same category by all models compatible with the selected reference set (necessary assignments), whereas one alternative is classified by some ambiguity in classes H or M (possible assignments). Similarly, 17 out of the 28 alternatives from class M are classified in the same category by all models derived from the selected reference set. However, 11 alternatives from class M are classified with ambiguity: three can be classified in H or M, five can be classified in M or L, whereas three actions can be assigned to any of the three categories (H, M, L). Finally, 20 necessary assignments are specified for alternatives of class L, whereas the remaining 17 alternatives of this class are assigned to categories M or L (possible assignments).

Table 2.1: Necessary and possible assignments for the non-reference alternatives

		Actual class			
		H	M	L	Total
$\mathscr{N}$	H	4	0	0	4
	M	0	17	0	17
	L	0	0	20	20
$\mathscr{P}$	{H, M}	1	3	0	4
	{M, L}	0	5	17	22
	{H, M, L}	0	3	0	3
	Total	5	28	37	70

To examine the robustness of the above results a resampling exercise is conducted. In particular, first a subsample of 20 alternatives is selected, at random, from the initial chosen reference set of 30 actions. Using this subsample as a new reference set, the necessary and possible assignments are computed for all of the 70 non-reference alternatives. A single AVF model is also constructed through formulation (2.4)–(2.10) and it is used to specify a single assignment for each one of the non-reference actions. The same experiment is repeated 100 times, each based on a different random subsample (new reference set) of 20 alternatives.

In each one of the above 100 tests, the best and worst assignments are identified for all non-reference alternatives. Table 2.2 presents the average frequencies with which each non-reference action is classified in the three categories. The results are reported in comparison to the necessary and possible assignments identified through the original reference set of 30 actions. Discrepancies between the results from the full reference set and the ones obtained from the 100 random tests are shown in bold.

For the alternatives necessarily assigned to category H, the simulation tests are mostly consistent with the necessary assignments. There is only a small likelihood (2.5 %) that an action necessarily assigned to class H under the full set might be

downgraded to category M if the reference set changes. However, the discrepancies for the two other categories are higher. For instance, for the alternatives that are

Table 2.2: Classification frequencies (in %) with the full set of AVFs corresponding to different perturbations of the reference set

		Best assignments			Worst assignments		
		H	M	L	H	M	L
$\mathscr{N}$	H	100.0	–	–	97.5	**2.5**	–
	M	**22.6**	77.4	–	–	80.3	**19.7**
	L	**4.6**	**25.6**	69.9	–	–	100.0
$\mathscr{P}$	{H, M}	100.0	–	–	–	84.8	**15.3**
	{M, L}	**8.0**	92.0	–	–	–	100.0
	{H, M, L}	100.0	–	–	–	–	100.0

necessarily assigned to category M with the full reference set, there is a significant likelihood (22.6 %) that will be upgraded to category H if the reference set changes. There is also a notable likelihood (19.7 %) for downgrading these alternatives to the low performance class L. Thus, claiming that these alternatives are consistently assigned to class M under all models compatible with the reference, does not seem to be a very robust conclusion, because variations of the reference set often lead to different outcomes.

The same also holds true for alternatives that are necessarily assigned to the low performance class L under the full reference set. In this case, there is notable likelihood (25.6 %) that they could be upgraded to the medium performance category M with a perturbed reference set, whereas the likelihood of an even further upgrade to class H is 4.6 %.

Similar discrepancies are also observed for the possible assignments, which are expressed in interval form. For instance, focusing on the alternatives that can be classified in H or M under the full reference set, the simulation test indicates that they could actually be classified to category L with some perturbation of the reference set.

Table 2.3 presents similar results with a single AVF model, obtained through the solution of problem (2.4)–(2.10) for each reference set in the 100 test runs. In this case smaller discrepancies are observed (shown in bold) between the results obtained with a single decision model (columns) and the necessary/possible assignments derived from the full reference set (rows). This should be of no surprise, as a single model does not provide information about extreme assignments like those considered in the above results.

The above obtained results support the argument in this study that similarly to point recommendations derived with a single decision model (AVF), interval results formulated on a set of decision models are also subject to the robustness concern when the reference data change.

Table 2.3: Classification frequencies (in %) with a single AVF for random perturbations of the reference set

		Model assignments		
		H	M	L
$\mathscr{N}$	H	99.8	**0.2**	–
	M	**2.8**	95.2	**2.1**
	L	**2.6**	**3.4**	94.0
$\mathscr{P}$	{H, M}	37.8	60.5	**1.8**
	{M, L}	**0.8**	32.3	66.9
	{H, M, L}	29.3	50.3	20.3

Table 2.4 reports some results about the support measure and the uncertainty of the assignments for the non-reference alternatives. Uncertainty is defined as the entropy of the assignments over the 100 test runs, with higher entropy values indicating higher ambiguity in the obtained classifications. Results are presented for the extreme (best and worst) assignments as well as for the assignments obtained with a single AVF. For the extreme assignments only the cases with positive support are considered because, as explained earlier a zero support indicates that the possible assignments cover all classes (e.g., from H to L in this example). For the results of the single AVF we also consider the cases with zero support to examine how ambiguous alternatives are classified when a single decision model is used. The obtained results clearly indicate that higher support is associated with lower ambiguity (i.e., lower entropy values) for all classifications, both the interval ones and the single AVF model assignments.

Table 2.4: Entropy of assignments vs support

Support	Best	Worst	Support	Single AVF
1	0.463	0.306	0	0.784
2	0.386	0.094	1	0.300
3	0.288	0.005	$\geq$2	0.212
4	0.080	0.006		
$\geq$5	0.007	0.004		

Regarding the two robustness indicators (2.16)–(2.18) that consider the size of the feasible polyhedron, they were found to be highly correlated to each other (Pearson correlation higher than 0.85) and strongly negatively correlated to the support measure (correlation about -0.6). The latter result implies the robustness of the assignments for non-reference alternatives with low support can be improved by reconsidering the evaluations of the supporting reference actions.

Table 2.5 provides details about the average values of the robustness indicators R and V, as defined by (2.16)–(2.18), for all assignments of the non-reference alter-

natives (the results are averages over the 100 tests). It is evident that both indicators attain their maximum values when the alternatives are classified in their respective necessary assignments. For instance, for alternatives assigned in category H by all models compatible with the full reference set, both R and V are equal to one for class H, whereas their value is lower for classes M ($R = 0.81$, $V = 0.76$) and L ($R = 0.84$, $V = 0.75$). Thus, both indicators confirm that H is the most robust assignment for these alternatives. The same holds for alternatives necessarily assigned to classes M and L using the full reference set. For alternatives for which the full reference set indicates that the can be classified in H or M (possible assignments), again the two indicators verify that these are the most robust conclusions (classes H and M correspond to higher values in R and V compared to class L). Similar, conclusions are also drawn for alternatives possibly assigned to M or L. These results, indicate that the two proposed robustness indicators are in accordance with the definitions of necessary and possible assignments, and enhance them with additional information that provides an analytic estimate of the robustness of the results, without requiring to resort to approximate simulation-based approaches.

Table 2.5: The robustness indicators for all assignment results (non-reference alternatives)

	R			V		
$\mathscr{N}/\mathscr{P}$	H	M	L	H	M	L
H	1.00	0.81	0.84	1.00	0.76	0.75
M	0.87	1.00	0.95	0.78	1.00	0.82
L	0.80	0.84	1.00	0.79	0.81	1.00
{H, M}	0.89	0.99	0.85	0.88	0.97	0.76
{M, L}	0.81	0.92	0.98	0.77	0.88	0.96
{H, M, L}	0.85	1.00	0.83	0.85	1.00	0.83

As a final test for the information content and validity of the two proposed indicators we consider the classification of the alternatives whose classification is ambiguous according to the reference set used in the analysis. These are 29 non-reference alternatives for which only their possible assignments could be defined (i.e., the alternatives classified in {H, M}, {M, L}, or {H, M, L}. To specify a single classification result for these cases we compare three different approaches:

1. For each of the 100 perturbations of the reference set, construct a single AVF model, use it to classify the alternatives, and finally use a majority rule to aggregate the 100 results for each alternative and specify the most appropriate class assignment.
2. Classify the alternatives to the class for which the R measure is highest.
3. Classify the alternatives to the class for which the V measure is highest.

The results of these three procedures are compared against the actual classification of the alternatives. The accuracy rate (i.e., the percentage of correct classifications)

for the assignments obtained through the majority rule was found to be 89.7 %, the assignments with the R measure had an accuracy rate of 82.8 %, whereas using the V measure led to an accuracy of 96.6 %. These results indicate that the two robustness indicators can constitute the basis for formulating good recommendations about the most appropriate classification when a reference set leads to ambiguous conclusions. Between the two indicators, the one based on the volume of the ellipsoid inscribed inside the feasible polyhedron (V) appears to provide better results.

2.5 Conclusions and Future Research

The robustness of MCDA models has been an active research topic recently having attracted a lot of interest from different perspectives. In this chapter we focused on the PDA framework for constructing decision models from data related to classification problems. PDA is based on a data-driven scheme. As such, changes in the data used to construct a decision model can have a significant impact on the results.

Motivated by this fact, this study presented simple, yet effective ways to assess the robustness of MCDA models in the form of AVFs for classification problems. The proposed measures provide analytic estimates of the ambiguity resulting from the information that a given data set provides, based on tools and techniques from optimization theory. The analytic form of the measures introduced in this study makes them applicable to all cases, even when dealing with large problem instances (i.e., reference sets with many actions and criteria).

The illustrative results presented in this chapter indicate that the proposed measures enhance existing robust MCDA techniques with additional information. Their connection with the concept of robustness in the data-driven context explained above was verified and their usefulness for formulating better decision recommendations was demonstrated.

However, the positive properties of the measures introduced in this study and the preliminary results should be further explored. To this end, applications to large real data sets and further experimental testing will provide further insights. Comparisons with simulation-based approaches could also be useful to construct an unified framework for analyzing robustness and assess the statistical properties of the proposed measures. Finally, extensions to other types of decision problems, including ordinal regression [11] should be examined, together with an analysis of cases where inconsistencies, uncertainties, and fuzziness are present in the data.

References

1. Bous, B., Fortemps, P., Glineur, F., Pirlot, M.: ACUTA: a novel method for eliciting additive value functions on the basis of holistic preference statements. Eur. J. Oper. Res. **206**(2), 435–444 (2010)

2. Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A., Vincke, P.: Evaluation and Decision Models with Multiple Criteria: Stepping Stones for the Analyst, vol. 86. Springer, Berlin (2006)
3. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge, MA (2004)
4. Doumpos, M., Zopounidis, C.: Multicriteria Decision Aid Classification Methods. Kluwer Academic Publishers, Dordrecht (2002)
5. Doumpos, M., Zopounidis, C.: Regularized estimation for preference disaggregation in multiple criteria decision making. Comput. Optim. Appl. **38**(1), 61–80 (2007)
6. Doumpos, M., Zopounidis, C.: Preference disaggregation and statistical learning for multicriteria decision support: a review. Eur. J. Oper. Res. **209**(3), 203–214 (2011)
7. Doumpos, M., Zopounidis, C.: The robustness concern in preference disaggregation approaches for decision aiding: an overview. In: Rassias, T., Floudas, C., Butenko, S. (eds.) Optimization in Science and Engineering, pp. 157–177. Springer, New York (2014)
8. Doumpos, M., Zopounidis, C., Galariotis, E.: Inferring robust decision models in multicriteria classification problems: an experimental analysis. Eur. J. Oper. Res. **236**(2), 601–611 (2014)
9. Greco, S., Matarazzo, B., Słowiński, R.: Rough sets theory for multicriteria decision analysis. Eur. J. Oper. Res. **129**(1), 1–47 (2001)
10. Greco, S., Mousseau, V., Słowiński, R.: Ordinal regression revisited: multiple criteria ranking using a set of additive value functions. Eur. J. Oper. Res. **191**(2), 416–436 (2008)
11. Greco, S., Słowiński, R., Figueira, J., Mousseau, V.: Robust ordinal regression. In: Ehrgott, M., Greco, S., Figueira, J. (eds.) Trends in Multiple Criteria Decision Analysis, pp. 241–283. Springer, New York (2008)
12. Greco, S., Mousseau, V., Słowiński, R.: Multiple criteria sorting with a set of additive value functions. Eur. J. Oper. Res. **207**(3), 1455–1470 (2010)
13. Greco, S., Kadziński, M., Słowiński, R.: Selection of a representative value function in robust multiple criteria sorting. Comput. Oper. Res. **38**(11), 1620–1637 (2011)
14. Jacquet-Lagrèze, E., Siskos, Y.: Assessing a set of additive utility functions for multicriteria decision making: the UTA method. Eur. J. Oper. Res. **10**, 151–164 (1982)
15. Jacquet-Lagrèze, E., Siskos, Y.: Preference disaggregation: 20 years of MCDA experience. Eur. J. Oper. Res. **130**, 233–245 (2001)
16. Kadziński, M., Tervonen, T.: Stochastic ordinal regression for multiple criteria sorting problems. Decis. Support Syst. **55**(1), 55–66 (2013)
17. Keeney, R., Raiffa, H.: Decisions with Multiple Objectives: Preferences and Value Trade-offs. Cambridge University Press, Cambridge (1993)
18. Mousseau, V., Slowinski, R., Zielniewicz, P.: A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. Comput. Oper. Res. **27**(7), 757–777 (2000)
19. Mousseau, V., Figueira, J., Dias, L., Gomes da Silva, C., Clímaco, J.: Resolving inconsistencies among constraints on the parameters of an MCDA model. Eur. J. Oper. Res. **147**(1), 72–93 (2003)
20. Roy, B.: The outranking approach and the foundations of ELECTRE methods. Theor. Decis. **31**, 49–73 (1991)
21. Roy, B.: Robustness in operational research and decision aiding: a multi-faceted issue. Eur. J. Oper. Res. **2010**(3), 629–638 (2010)
22. Siskos, Y., Grigoroudis, E.: New trends in aggregation-disaggregation approaches. In: Zopounidis, C., Pardalos, P. (eds.) Handbook of Multicriteria Analysis, pp. 189–214. Springer, Berlin, Heidelberg (2010)
23. Tervonen, T., Figueira, J., Lahdelma, R., Dias, J., Salminen, P.: A stochastic method for robustness analysis in sorting problems. Eur. J. Oper. Res. **192**(1), 236–242 (2009)
24. Vetschera, R., Chen, Y., Hipel, K., Marc Kilgour, D.: Robustness and information levels in case-based multiple criteria sorting. Eur. J. Oper. Res. **202**(3), 841–852 (2010)

25. Vincke, P.: Multicriteria Decision Aid. Wiley, New York (1992)
26. Zopounidis, C., Doumpos, M.: Building additive utilities for multi-group hierarchical discrimination: the M.H.Dis method. Optim. Method Softw. **14**(3), 219–240 (2000)
27. Zopounidis, C., Doumpos, M.: Multicriteria classification and sorting methods: a literature review. Eur. J. Oper. Res. **138**(2), 229–246 (2002)

Chapter 3
Robustness for Adversarial Risk Analysis

David Ríos Insua, Fabrizio Ruggeri, Cesar Alfaro, and Javier Gomez

Abstract Adversarial Risk Analysis is an emergent paradigm for supporting a decision maker who faces adversaries in problems in which the consequences are random and depend on the actions of all participating agents. In this chapter, we outline a framework for robust analysis methods in Adversarial Risk Analysis. Our discussion focuses on security applications.

3.1 Introduction

Large scale terrorist events like S-11 led to huge security investments. In turn, this has promoted many modeling efforts to support how to efficiently allocate such resources. Parnell et al. [15] provided an in-depth review for the US National Academy of Sciences on bio-terrorism assessment, concluding, among other things, that traditional risk analysis tools, like event trees, are not adequate in this application area for not accounting for adversarial intentionality; the critical and, in many contexts, doubtful common knowledge assumptions of game theoretic approaches;

D. Ríos Insua (✉)
Instituto de Ciencias Matemáticas, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain
e-mail: david.rios@icmat.es

F. Ruggeri
Consiglio Nazionale delle Ricerche, Istituto di Matematica Applicata e Tecnologie Informatiche, Via Bassini 15, 20133 Milano, Italy
e-mail: fabrizio@mi.imati.cnr.it

C. Alfaro • J. Gomez
Department of Statistics and Operations Research, Rey Juan Carlos University, Fuenlabrada, 28943 Madrid, Spain
e-mail: cesar.alfaro@urjc.es; javier.gomez@urjc.es

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_3

and, finally, the problems of decision analytic based approaches in forecasting adversarial actions. Merrick and Parnell [12] reviewed numerous approaches in this research area, commenting favorably on Adversarial Risk Analysis (ARA), which is a framework to manage risks derived from actions of intelligent adversaries, see [20] or [1].

ARA aims at providing one-sided prescriptive support to one of the intervening agents, the Defender (D, she), based on a subjective expected utility model treating the adversary's decisions as uncertainties. To do so, we model the adversary's (A, Attacker, he) decision making problem and, assuming that he is an expected utility maximizer, try to assess his probabilities and utilities. We can consequently forecast his optimal action. However, our uncertainty about the adversary's probabilities and utilities is propagated to his decision, leading to a random optimal adversary decision which provides us with the required distribution over the Attacker's decision. Sometimes such assessments may lead to a hierarchy of nested decision problems, as described in [17], similar to the concept of level-k thinking, see [24]. In contrast with game theoretic approaches, we do not assume the standard, but unrealistic, common knowledge hypothesis, see [5], according to which the agents share information about their utilities and probabilities.

A critical issue in ARA is elicitation. As in any subjective Bayesian analysis, one needs personal probabilities over the parameters in the problem. Obtaining them is not easy and we need to cope with many biases, see e.g., [14]. This is aggravated in our context because of the involved strategic considerations. Nau [13] as well as Wang and Bier [26] provide discussions of elicitation in the context of adversarial situations.

The practical difficulty of elicitation raises the question of robustness. One wants an ARA to be robust to the elicited probabilities and utilities, the model entertained and, when available, the data. A good way forward is sensitivity analysis. The above mentioned review by Parnell [15] recommends it, and Von Winterfeldt and O'Sullivan [25] perform a systematic sensitivity analysis with respect to elicited probabilities in an event tree concerning MANPADS. A different approach is taken by Kardes [9], who considers robust stochastic games.

Robust Bayesian analysis facilitates finding the entire set of posterior distributions for a parameter when the prior lies within a class of distributions. The results are typically expressed in terms of upper and lower bounds on probabilities and expected utilities. Berger et al. [2] review this methodology which has yet to be used in ARA. The only direct application is given by McLay et al. [11], who point the way towards a principled means to incorporate robustness into ARA. They consider a level-k thinking analysis of the sequential Defend-Attack game in which the Attacker imperfectly observes the decision made by the Defender. The game is modeled through an information structure comprising several signals and, conditional on the defense choice, there is a specified distribution over the signals, a model initially proposed by Rothschild et al. [22]. Robustification occurs by setting upper and lower bounds over parameters for which distributions must be elicited, and then calculating the outcome under the worst case combination of upper and lower values.

This chapter provides a complete outline of the role of robust methods in ARA. After introducing basic notions in Bayesian robustness, we first describe the robust ARA approach for sequential games and, then, for simultaneous games. In both cases, we start by computing the game theoretic solution. We apply robust concepts to assess such solution. If it is not robust, we use the ARA approach to find an alternative solution. Again, we criticize it through robust ideas. If the solution is still unstable, we may appeal to conventional robust concepts, such as the γ maximin. We illustrate the ideas with a simple numerical example concerning routing security.

3.2 Bayesian Robustness

We present here the basic ideas on Bayesian robustness. We refer to Ríos Insua and Ruggeri [18] for an in-depth overview. In the Bayesian approach to inference, prediction and decision making, the interest frequently lies on the behavior of the posterior distribution on a parameter θ obtained by combining experimental evidence provided by the likelihood and expert knowledge expressed through the prior distribution, via Bayes theorem. This is used to compute posterior (and predictive) expectations of functions $g(\theta)$ which typically will be set indicators, powers or utility functions, providing, respectively, set probabilities, moments and expected utilities. The robust Bayesian approach stems from the practical difficulty of specifying a unique prior distribution and/or a unique utility function, corresponding, respectively, to the expert's beliefs and the decision maker's preferences. Therefore, classes of priors and/or utilities are entertained and the consequences of different possible choices of such pairs are evaluated through synthetic indices which determine whether the quantity of interest is subject to small or large variations when changing the prior/utility, i.e. whether there is robustness or not.

In accordance with the content of this chapter, we shall consider utilities u in a class $\mathscr{U}$ and probability measures p in a class $\mathscr{P}$ (without distinguishing whether they are priors or posteriors). We suppose that the probability measure p has a density $p(s)$ over the states s, and the utility function has the form $u(d,s)$, where d is an action (decision) in the feasible set $\mathscr{D}$. We are interested in computing the expected utilities $\psi_{up}(d) = \int u(d,s)p(s)ds$ for various alternatives d and the feasible alternative $d^*_{up} \in \mathscr{D}$ maximizing expected utility, given such choice u and p.

In a robust context, the interest would typically be in the ranges that relevant quantities span when p and u vary in the class, e.g. the range of the expected utility for a certain alternative d

$$\rho_\psi(d) = \sup_{p\in\mathscr{P},u\in\mathscr{U}} \psi_{up}(d) - \inf_{p\in\mathscr{P},u\in\mathscr{U}} \psi_{up}(d),$$

or the distance between the optimal alternative and a reference alternative d^*

$$\rho_d = \sup_{p\in\mathscr{P},u\in\mathscr{U}} e(d^*_{up}, d^*),$$

for some distance e. Looking at ρ_d, we claim that there is robustness if its value is *small* with respect to the entertained problem and the decision maker's perception. In this case, essentially any p and u, and the corresponding d^*_{up}, may be used for decision making purposes. Otherwise, efforts are required to get smaller classes until either robustness can be achieved or no further refinement is possible.

In the latter case, some criterion could be introduced to choose a pair (p,u) and the corresponding d^*_{up}. A possible choice for a decision could be the minimum regret decision,

$$\hat{d} = \text{argmin}_{d\in\mathscr{D}} \max_{p\in\mathscr{P},u\in\mathscr{U}} \left[\psi_{up}(d^*_{up}) - \psi_{up}(d)\right].$$

For a related discussion see [19]. In particular, the decision $\hat{d}$ is conservative in the sense that it protects against the worst loss in expected utility when replacing an optimal decision d^*_{up} by another one.

3.3 Sequential Games

We start by considering sequential games: one agent first makes her decision and, then, the other agent implements his alternative. As an example, imagine a case in which a company deploys their cybersecurity countermeasures and then, observing them, a hacker decides whether he launches an attack or not towards such company.

Specifically, we consider a Defend-Attack situation in which a Defender chooses a defense $d \in \mathscr{D}$ and, then, the Attacker, having observed the defense, chooses his attack $a \in \mathscr{A}$. The corresponding bi-agent influence diagram is shown in Fig. 3.1. An arc reflects that the Defender's choice is observed by the Attacker. The consequences for both players depend on the success s of the attack. Each decision maker assesses differently the probability of the result of an attack, which depends on the defense and attack adopted: $p_D(s \mid d,a)$ and $p_A(s \mid d,a)$. The utility function of the Defender $u_D(d,s)$ depends on her chosen defense and the result of the attack. Similarly, the Attacker's utility function is $u_A(a,s)$. We first recall the standard game theoretic approach and check its robustness. We then present the ARA solution and, again, provide a robust analysis.

3.3.1 Game Theoretic Solution and Robustness

The standard game theoretic solution does not require the Attacker to know the Defender's probabilities and utilities, since he observes the Defender's actions. However, the Defender needs to know the Attacker's utilities and probabilities (u_A, p_A), an example of common knowledge. We then proceed as follows. First, we compute the expected utilities of the players at node S in Fig. 3.1:

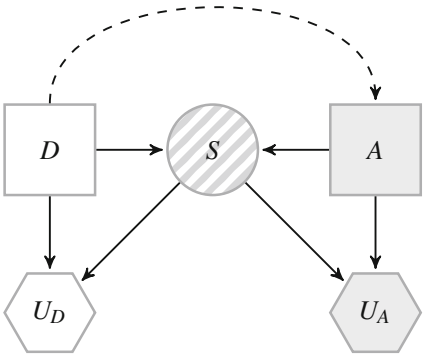


Fig. 3.1: The two player sequential decision game

$$\psi_A(a,d) = \int u_A(a,s)p_A(s|a,d)ds, \tag{3.1}$$

$$\psi_D(a,d) = \int u_D(d,s)p_D(s|a,d)ds.$$

Then, we compute the Attacker's best response to the Defender's action d, which is

$$a^*(d) = \text{argmax}_{a\in\mathscr{A}}\psi_A(a,d).$$

Knowing this, the Defender's optimal action is, then,

$$d^*_{GT} = \text{argmax}_{d\in\mathscr{D}}\psi_D(a^*(d),d).$$

The solution $\left(a^*(d^*_{GT}), d^*_{GT}\right)$ is a Nash equilibrium and, indeed, a sub-game perfect equilibrium, see [5]. We call d^*_{GT} the Nash defense.

3.3.1.1 Robustness of the Game Theoretic Solution

Since we are supporting the Defender, we could argue that we know reasonably well (u_D, p_D). However, we would contend that knowledge about (u_A, p_A) is that precise, since it would require the Attacker to reveal them (common knowledge). This is questionable in many application areas including security, cybersecurity and competitive marketing. We may use robust methods to criticize such information and, consequently, assess the game theoretic solution.

As discussed in Sect. 3.2, from a conceptual point of view, to perform robustness we may consider classes for the Attacker's utilities and probabilities that we model through $u \in \mathscr{U}_A$, $p \in \mathscr{P}_A$. Then, mimicking the approach above, for each feasible (u,p) we could:

- Compute the expected utilities $(\psi_A^{u,p}(d,a), \psi_D^{u,p}(d,a))$ at node S in Fig. 3.1.
- Compute the best response attack $a^*_{u,p}(d)$ for each d.
- Compute the optimal defense $d^*_{u,p}$.

Then, if $d^*_{u,p}$ remains reasonably stable for the allowed perturbations of u and p, with $u \in \mathscr{U}_A$, $p \in \mathscr{P}_A$, the game theoretic solution seems robust. However, if $d^*_{u,p}$ is not that stable, we have an issue which questions, at first sight, the relevance of the proposed Nash defense d^*_{GT}. At a deeper level, it also questions the appropriateness of the (u_A, p_A) assessment, actually serving to criticize the game theoretic assumptions, specially that of common knowledge, see [16] or [10].

From an operational point of view, the above robustness analysis scheme for the game theoretic approach boils down to two computational issues:

- Exploring the whole range of perturbations $u \in \mathscr{U}_A$, $p \in \mathscr{P}_A$. In some cases, for classes of probabilities and utilities widely studied in the robust Bayesian literature, see [2], it is possible to identify the extremal elements of $\mathscr{U}_A$ and $\mathscr{P}_A$ and compute upper and lower bounds on the quantities of interest (namely optimal decisions $d^*_{u,p}$ and their expected utilities), through numerical optimization methods. Another possible approach would be to randomly sample elements u, p from the sets $\mathscr{U}_A, \mathscr{P}_A$ and check for eventual large variations in $d^*_{u,p}$ (and their expected utilities).
- Declaring whether the effects induced by changes over $d^*_{u,p}$ and the expected utility are sufficiently small. As discussed in Sect. 3.2, a possible criterion could be given by the range spanned by $d^*_{u,p}$ as utility and probability vary in the classes, i.e. $u \in \mathscr{U}_A$ and $p \in \mathscr{P}_A$, respectively. Regarding the effects on the expected utility, a criterion of interest could be based on the regret $r_{u,p}(d^*_{GT})$ given by the difference in expected utility when considering, for a given pair (u, p), the Nash defense d^*_{GT} and the optimal defense $d^*_{u,p}$. A small value of $\sup_{(u,p)\in\mathscr{U}_A\times\mathscr{P}_A} r_{u,p}(d^*_{GT})$ would denote robustness with respect to the choice of utility and probability and, therefore, any pair (u, p) can be chosen as opinion on the Attacker's behavior with no significant change in the consequences. If robustness is not achieved, then we could undertake a minimum regret approach as discussed in Sect. 3.2.

An alternative would be to move to ARA, as discussed next.

3.3.2 ARA Solution and Robustness

If the game theoretic solution is not robust, then we need to address the issue. One way forward is to perform an ARA approach. For this, we weaken the common knowledge assumption. In the sequential game, this means that the Defender does not know (p_A, u_A). The problem she faces is depicted in Fig. 3.2.

To solve her problem, the Defender requires more information than $p_D(s|a,d)$ and $u_D(d,s)$, available from our earlier discussion. She also needs $p_D(a|d)$, which is her assessment of the probability that the Attacker will choose attack a after having observed that she has chosen the defense d. Once the Defender has completed these assessments, she can solve the problem. Indeed, the expected utility of d would be

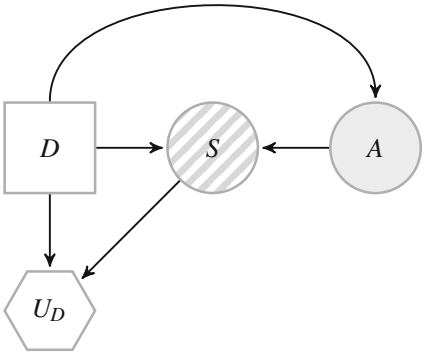


Fig. 3.2: The decision problem as seen by Defender

$$\psi_D(d) = \int \psi_D(a,d) p_D(a|d) da = \int \left[\int u_D(d,s) p_D(s|a,d) ds \right] p_D(a|d) da.$$

Finally, her optimal decision would be $d^*_{ARA} = \text{argmax}_{d \in \mathscr{D}} \psi_D(d)$. Note that, in terms of classic game theory, the solution d^*_{ARA} for our sequential game may not correspond to a Nash equilibrium, see the example in Sect. 3.5.

Eliciting $p_D(a|d)$ requires the Defender to analyze the problem from the Attacker's perspective.

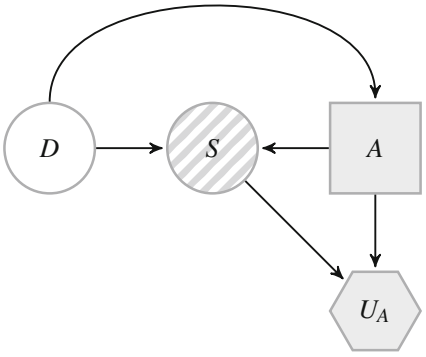


Fig. 3.3: Defender's analysis of Attacker's problem

First, the Defender puts herself in the Attacker's shoes, and thinks about his decision problem. Figure 3.3 represents the Attacker's problem, as seen by the Defender, assuming he is an expected utility maximizer. The Defender will use all the information and judgment that she can obtain about the Attacker's utilities and probabilities. Instead of using point estimates for p_A and u_A to find the Attacker's optimal decision $a^*(d)$ for a given d, the Defender's uncertainty about the Attacker's decision should derive from her uncertainty about the Attacker's (p_A, u_A), through a distribution F

on the space of utilities and probabilities, which we designate random probabilities and utilities. This induces a distribution over the Attacker's expected utility in (3.1), where the random expected utility for A would be

$$\Psi_A(a,d) = \int U_A(a,s)P_A(s|a,d)ds,$$

for $(P_A, U_A) \sim F$. Then, the Defender would find

$$p_D(a|d) = \mathbb{P}_F[a = \text{argmax}_{x\in\mathscr{A}}\Psi_A(x,d)],$$

in the discrete case and, similarly, in the continuous case. We can use Monte Carlo simulation to approximate $p_D(a|d)$ by drawing N samples $\left\{\left(P_A^i, U_A^i\right)\right\}_{i=1}^N$ from F and setting

$$\hat{p}_D(a|d) \approx \frac{\#\{a = \text{argmax}_{x\in\mathscr{A}}\ \Psi_A^i(x,d)\}}{N}, \tag{3.2}$$

where $\Psi_A^i(a,d) = \int U_A^i(a,s)P_A^i(s|a,d)ds$.

3.3.2.1 Robust Analysis

The above approach leads to a Bayesian decision analysis problem with the peculiarity that we have a complex procedure to forecast the adversarial actions. To do so, we formulate the adversary decision making problem and propagate our uncertainty about the adversary judgments to the optimal adversarial action.

We could then think about performing a robust Bayesian analysis. The inputs to the Defender's decision analysis are $(u_D(d,s), p_D(s|a,d), p_D(a|d))$. We focus here on sensitivity to the last component $p_D(a|d)$, surely the most contentious one, attained through adversarial calculations based on the proposed $U_A(a,s), P_A(s|a,d)$. For that, we define classes $\mathscr{U}_A$, $\mathscr{P}_A$ of random utilities and probabilities. For each pair U,P in such class, we define $p_D^{UP}(a|d)$ through the ARA approach which, in turn, leads to d_{ARA}^{*UP}.

Then, it is possible to consider the impact of the imprecision about U and P over three quantities: $p_D^{UP}(a|d)$, d_{ARA}^{*UP} and $\psi(d_{ARA}^{*UP})$. The first quantity requires the comparison of densities (actually of their Monte Carlo approximations) using indices like the Kullback-Leibler divergence or Gini index. For the first and second quantities, the interest centers around the variation of the decision (for the Defender), whereas for the third one, the focus is on the expected utility of the decision. The last quantity should be of major interest. In all three cases, we say that robustness holds when the value of interest does not change much, whereas additional analysis should be taken otherwise, as described in Sect. 3.2. In particular, if the distributions $p_D^{UP}(a|d)$ do not differ too much, it is possible to choose one of them and use d_{ARA}^{*UP} directly.

3.3.3 A Full Robust Solution

If the ARA analysis is not robust, we may opt for gathering additional information to reduce the classes $\mathscr{U}_A$ and $\mathscr{P}_A$. The choice of increasing the sample size in the Monte Carlo estimation $\hat{p}_D(a|d)$ in (3.2) would be useful in reducing the variability of the distribution. However, it will typically be ineffective in increasing robustness.

Once all possible sources of information have been exploited to try to increase robustness about d^*_{ARA} and $\psi(d^*_{ARA})$, then some extra criterion has to be introduced to make a decision and report a value about the quantity of interest. In any case, such decision should be reported with the warning of lack of robustness. As discussed in Sect. 3.2, we could consider the decision d^*_R minimizing the maximum regret, i.e.

$$\min_d \max_{U\in U_A, P\in P_A} \left[\int \psi_D(a, d^{*UP}_{ARA}) p^{UP}_D(a|d^{*UP}_{ARA}) da - \int \psi_D(a,d) p^{UP}_D(a|d) da \right].$$

3.4 Simultaneous Games

We discuss now the simultaneous game model: two agents choose their decisions, without knowing the action selected by each other. Among others, see [27] for a related discussion within a game theoretic framework. As an example, imagine a case in which the EASA decides whether to introduce undercover marshals in an airplane that might, or not, be hijacked by terrorists.

Assume that the adversaries have alternative sets $\mathscr{D}$ and $\mathscr{A}$ of defenses and attacks, respectively. The only relevant uncertainty is S, denoting the success s of the attack. Each decision maker assesses differently the probability of the result of the attack, which depends on the defense and attack adopted: $p_D(s \mid d,a)$ and $p_A(s \mid d,a)$. The utility function of the Defender $u_D(d,s)$ depends on her chosen defense and the result of the attack. Similarly, the Attacker's utility function is $u_A(a,s)$, as illustrated in Fig. 3.4.

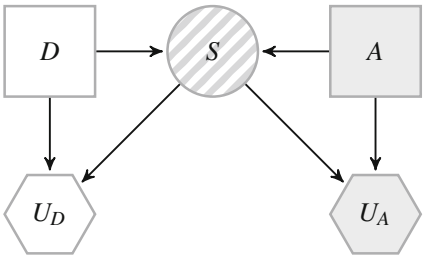


Fig. 3.4: BAID for the simultaneous Defend-Attack model

3.4.1 Game Theoretic Solution

Under common knowledge, preferences and beliefs from both the Defender and the Attacker, (u_D, p_D) and (u_A, p_A) respectively, are disclosed. Therefore, each adversary knows the expected utility that each pair $(d,a) \in \mathscr{D} \times \mathscr{A}$ would provide to both of them, computed through

$$\psi_D(d,a) = \int u_D(d,s) p_D(s|a,d) ds,$$

$$\psi_A(d,a) = \int u_A(a,s) p_A(s|a,d) ds.$$

A Nash equilibrium (d^*_{GT}, a^*_{GT}) for this game would satisfy

$$\psi_D(d^*_{GT}, a^*_{GT}) \geq \psi_D(d, a^*_{GT}) \ \forall d \in \mathscr{D} \quad \text{and}$$
$$\psi_A(d^*_{GT}, a^*_{GT}) \geq \psi_A(d^*_{GT}, a) \ \forall a \in \mathscr{A}.$$

Finding Nash equilibria may require the use of randomized strategies, see [4]. There could be several equilibria, with no unambiguous criteria to further discern among them, see [16] for a discussion.

If utilities and probabilities are not common knowledge among the adversaries, a game-theoretic approach proceeds by modeling the game as one with incomplete information, see [6–8], by introducing the notion of player types. Each player will be of a certain type which is known to him but not to his opponent: a player's type represents the private information he may have. Harsanyi proposes the Bayes-Nash equilibrium as a solution concept, still under a strong common knowledge assumption: the adversaries' beliefs about the opponent's types are common knowledge and modeled through a common prior distribution. Moreover, it is assumed that the players' beliefs about other uncertainties in the problem are also common knowledge. Again randomized strategies might be required to find such equilibria.

3.4.1.1 Robustness of the Game Theoretic Solution

We could argue that we know reasonably well (u_D, p_D), since we are supporting the Defender. However, we would contend that (u_A, p_A) is properly known, since it requires common knowledge, which is questionable. To address this concern, we perform a robust analysis of the Defender's decision at the Nash equilibrium.

For that, we would consider classes for the Attacker's utilities and probabilities represented as $u \in \mathscr{U}_A$, $p \in \mathscr{P}_A$. Then, for each feasible (u,p) we could compute the corresponding Nash equilibrium (d^*_{up}, a^*_{up}). If d^*_{up} remains stable for the feasible perturbations of u and p, the game theoretic solution d^*_{GT} seems robust, from the perspective of the Defender. However, if d^*_{up} changes, specially the corresponding expected utility, we have a problem which questions, at first sight, the relevance of the proposed d^*_{GT} and, at a deeper level, the appropriateness of the (u_A, p_A) assessment,

actually serving to criticize the game theoretic approach at large and, in particular, the common knowledge assumption. The two computational issues about finding all possible optimal decisions and assessing robustness are dealt with as mentioned in Sect. 3.3.3.

Note that we could actually study robustness with respect to (u_D, p_D, u_A, p_A) and consider changes in $d^*_{u_A,p_A,u_D,p_D}$. In this case, if the Defender's Nash equilibrium decision is sensitive, we might question the Defender's knowledge, besides the game theory postulates.

3.4.2 ARA Solution and Robustness

If the Nash equilibrium is unstable, we may try an ARA approach. We have to weaken the common (prior) knowledge assumptions. As reflected in Fig. 3.5, the Defender has to choose a defense $d \in \mathscr{D}$, whose consequences depend on the success of an attack $a \in \mathscr{A}$ simultaneously chosen by the Attacker, which is, therefore, uncertain for the Defender at the time she makes her decision.

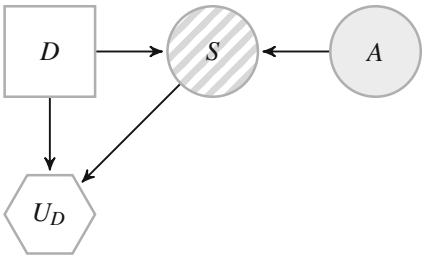


Fig. 3.5: The Defender's decision analysis

By standard Decision Theory, the Defender should maximize her expected utility, see [3]. The Defender knows her utility function $u_D(d,s)$ and her probability assessment p_D over S, conditional on (d,a). However, she does not know the Attacker's decision a at node A. She expresses her uncertainty through a probability distribution $p_D(a)$. Then, the optimization problem she should solve to find d^*_{ARA} is

$$\begin{aligned} \max_d \int \psi_D(a,d) p_D(a) da &= \max_d \int [\int u_D(d,s) p_D(s|a,d)\, ds]\, p_D(a)\, da \\ &= \max_d \int \int u_D(d,s) p_D(s|a,d) p_D(a)\, ds\, da. \end{aligned} \tag{3.3}$$

We could then perform a robust analysis based on $u_D, p_D(s|a,d)$ and $p_D(a)$. However, eliciting this last probability distribution is more difficult. We may use ARA as follows to get it.

Suppose the Defender thinks that the Attacker is an expected utility maximizer who tries to solve the decision problem shown in Fig. 3.6. The Attacker would look for the attack $a \in \mathscr{A}$ providing him maximum expected utility:

$$a^* = \arg\max_{a \in \mathscr{A}} \int \int u_A(a,s) p_A(s|a) p_A(d) ds\, dd.$$

In general, the Defender will be uncertain about the Attacker's utility function and probabilities, and she would consider random utilities and probabilities through $F = (U_A(a,s), P_A(s|a), P_A(d))$ and compute the random optimal alternative

$$A^*|D = \arg\max_{a \in \mathscr{A}} \int \int U_A(a,s) P_A(s|a) P_A(d) ds\, dd. \tag{3.4}$$

Then, we would make

$$p_D(a) = P(A^* = a|D)$$

in the discrete case and, similarly, in the continuous case.

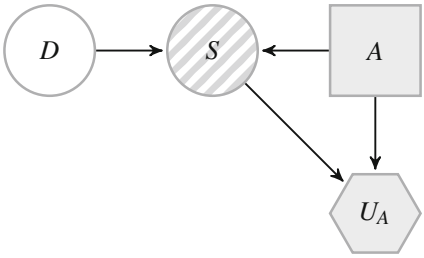


Fig. 3.6: The Attacker's decision analysis, as seen by the Defender

Note that $(U_A(a,s), P_A(s|a))$ would be comparatively easily elicited from the Defender, see examples in [1]. However, the elicitation of $P_A(d)$ may require further analysis leading to a next level of recursive thinking: the Defender would need to think about how the Attacker analyzes her problem. This is why we condition in (3.4) by (the distribution of) D.

In the above, the Defender presumes that the Attacker thinks she is an expected utility maximizer trying to solve a decision problem like that described in Fig. 3.5. Therefore, in order for the Defender to assess the required distribution, she will elicit (U_A, P_A) from her viewpoint, and assess $P_A(D)$ through the analysis of her decision problem, as thought by the Attacker, mimicking the resolution of problem (3.3) from the Attacker's perspective. This reduces the assessment of $P_A(D)$ to computing the distribution

$$D \mid A^1 \sim \arg\max_{d \in \mathscr{D}} \int \int U_D(d,s) P_D(S = s \mid d,a) P_D(A^1 = a) ds\, da,$$

assuming that the Defender is able to assess $P_D(A^1)$. A^1 represents the Attacker's decision within the Defender's second level of recursive thinking in the nested decision model used by the Defender to predict the Attacker's analysis of her decision problem. To complete the assessment, the Defender should elicit $(U_D, P_D) \sim G$, representing her probabilistic knowledge about how the Attacker may estimate her

utility function $u_D(d,s)$ and her probability p_D over $S|d,a$, when she analyzes how the Attacker thinks about her decision problem. The elicitation of $P_D(A^1)$ might require further recursive thinking from the Defender, see our final discussion.

3.4.2.1 Robustness

Performing a robust analysis for the ARA approach to the simultaneous game would be similar to what described earlier. Consider a class for $(U_A(a,s),P_A(s|a),P_A(d)) \in (\mathscr{U}_A,\mathscr{P}_A,\mathscr{Q}_A)$. We use (U,P,Q) to simplify the notation describing the elements in the classes. Then, for (U,P,Q) satisfying the constraints, replicating the approach above we could compute p_D^{UPQ} and d_{ARA}^{*UPQ}. If d_{ARA}^{*UPQ} remains stable with respect to changes in (U,P,Q), then the problem seems robust and we could apply the ARA approach with little concern. Otherwise, we could still use a robust solution concept, like the minimum regret mentioned in Sect. 3.2.

3.5 An Example

As an illustration, we consider a sequential defend-attack security routing problem. An organization needs to make a trip, either through a safe, but costly, route, or through a cheaper, but more dangerous, route. In this case, they may invest in security, rendering the route less dangerous. See [23] for a case concerning piracy in Somalia. Table 3.1 displays the consequences, expressed as costs, for various defend and attack possibilities.

Table 3.1: Loss function in routing problem

Defense	Attack	Attack result	Def. cons.	Att. cons.
Dang. prot	Attack	θ_1	$c\theta_1+K$	$-d\theta_1+B$
	No Attack		K	0
Dang. unp	Attack	θ_2	$c\theta_2$	$-d\theta_2+B$
	No attack		0	0
Safe			H	0

The following parameters are used:

- θ_1 represents the fraction of assets lost by the organization when attacked but protected.
- θ_2 represents the fraction of assets lost by the organization when attacked and not protected.
- c is the cost per unit of assets.
- K are the protection costs.
- H is the cost of going through the expensive route.

- d is the Attacker's gain per unit of assets lost by the Defender.
- B is the cost of an attack.

The Defender has beliefs for θ_i, with $\theta_i \sim \beta(a_i, b_i), i = 1, 2$. She is risk averse and her utility function is strategically equivalent to $-\exp(hx)$, where x is her cost and $h > 0$ is her risk aversion coefficient. The Attacker has different beliefs for θ_i with $\theta_i \sim \beta(c_i, e_i), i = 1, 2$. He is risk prone and his utility function is strategically equivalent to $\exp(-mx)$, where x is his cost and $m > 0$ is his risk proneness coefficient. Both agents expect θ_1 to be smaller than θ_2, but not necessarily. This may be reflected in the choice of the beta parameters, for example with $a_1/(a_1 + b_1) < a_2/(a_2 + b_2)$, in the case of the Defender. Table 3.2 provides the expected utilities for both agents under various interaction scenarios.

Table 3.2: Expected utilities in routing problem

Interaction	Eu. def	Eu. att
Prot, Att.	$-\int e^{h(c\theta_1+K)} f(\theta_1 \mid a_1, b_1) d\theta_1$	$\int e^{m(d\theta_1-B)} f(\theta_1 \mid c_1, e_1) d\theta_1$
Prot.,NoAtt.	$-e^{hK}$	1
NoProt.,Att.	$-\int e^{h(c\theta_2)} f(\theta_2 \mid a_2, b_2) d\theta_2$	$\int e^{m(d\theta_2-B)} f(\theta_2 \mid c_2, e_2) d\theta_2$
NoProt.,NoAtt.	-1	1
Safe	$-e^{hH}$	1

The problem may be viewed through the game tree in Fig. 3.7, where d_1 means going through the dangerous route but protected; d_2 means going through the dangerous route but unprotected; and, finally, d_3 means going through the safe route, whereas a means *attack* and $\bar{a}$ means *no attack*.

We are supporting the Defender and assess from her the values $c = 200{,}000$, $K = 50{,}000$, $H = 100{,}000$, $h = 3$. We also elicit from her the distributions $\beta(a_1, b_1)$, with mean 0.3 and standard deviation 0.07, leading to $a_1 = 12.325$, $b_1 = 28.76$; and $\beta(a_2, b_2)$, with mean 0.7 and standard deviation 0.18, leading to $a_2 = 3.815$, $b_2 = 1.635$.

3.5.1 Game Theoretic Approach

Under common knowledge, we assume the Defender knows that $d = 30{,}000$, $B = 10{,}000$, $m = 5$ and the distributions $\beta(c_1, e_1)$, with mean 0.313 and standard deviation 0.16, leading to $c_1 = 2.272$, $e_1 = 4.978$; and $\beta(c_2, e_2)$, with mean 0.324 and standard deviation 0.11, leading to $c_2 = 5.49$, $e_2 = 11.45$. We, then, proceed as follows:

- At node A_1, compute $\max(\psi_A(d_1, a), \psi_A(d_1, \bar{a}))$ and call the optimal action $a^*(d_1)$. In the example, we have $\max(1.001, 1) = 1.001$ and the optimal decision for the Attacker is a.

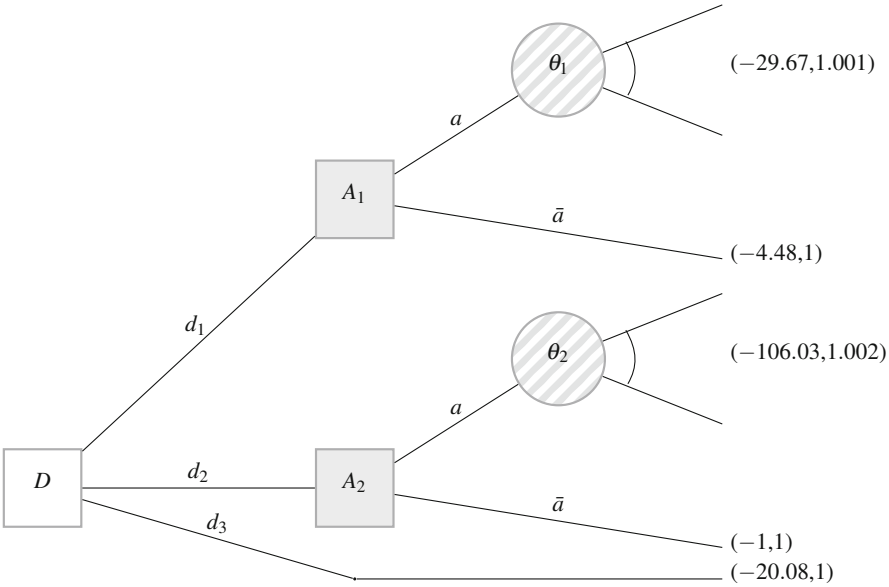


Fig. 3.7: Game tree for the routing problem

- At node A_2, compute max $(\psi_A(d_2,a),\psi_A(d_2,\bar{a}))$ and call the corresponding action $a^*(d_2)$. We have max $(1.002,1) = 1.002$ and the optimal decision for the Attacker is a.
- At node D, compute max $(\psi_D(d_1,a^*(d_1)),\psi_D(d_2,a^*(d_2)),\psi_D(d_3))$ and call the optimal action d^*_{GT}. In our case, max $(-29.67,-106.03,-20.08) = -20.08$ and the Nash defense d^*_{GT} is d_3, that is, to choose the safe route.

3.5.2 Robustness of the Game Theoretic Solution

We consider now the robustness of the game theoretic solution. We simplify and assume that the attack cost $B = 10{,}000$ is reasonably well known. Assume that d is not that well known and we express this through a constraint $d \in [10000, 50000]$. Similarly, suppose that $c_1 \in [0,3]$, $e_1 \in [1,6]$, $c_2 \in [2,8]$ and $e_2 \in [10,14]$. We sample randomly from these intervals 1000 times and repeat the procedure in Sect. 3.5.1.

The three defenses may be Nash, given the constraints. Indeed, based on the above sampling scheme, we estimate that the probabilities of the three alternatives being Nash are, respectively, 0.454, 0.236 and 0.31, therefore with no clear winner. The maximum loss when we implement the defense $d^*_{GT} = d_3$ is 19.08. This is deemed large enough and we need to perform an ARA approach.

3.5.3 ARA Approach

The problem faced by the Defender is described in the decision tree in Fig. 3.8.

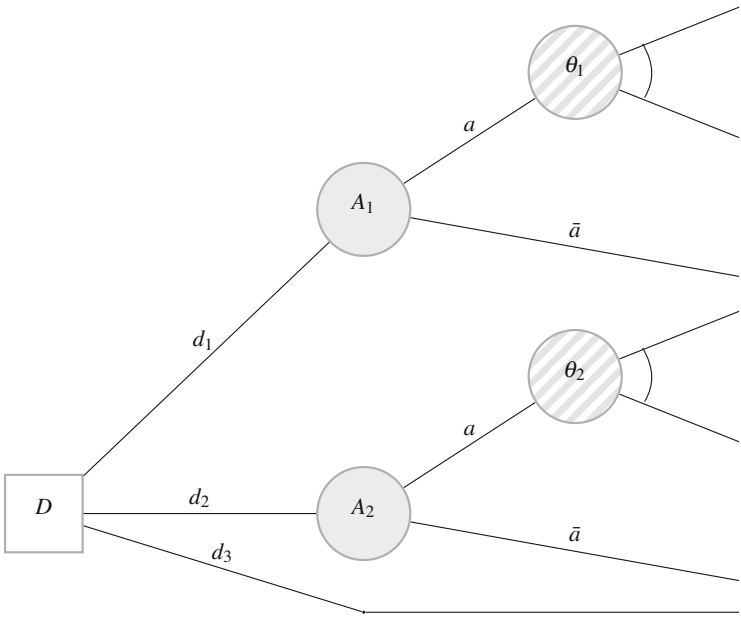


Fig. 3.8: Decision tree for the Defender in the routing problem

The expected utilities of the first two alternatives have the form

$$\psi_D(d_i) = p_D(a|d_i)\psi_D(d_i,a) + p_D(\bar{a}|d_i)\psi_D(d_i,\bar{a}), i = 1,2.$$

Thus, we need to assess the attack probabilities $p(a|d_i)$ given the implemented defense d_i.

We illustrate the estimation of $p_D(a|d_1)$. We assume that d, c_1, e_1, c_2, e_2 are uniformly distributed over the intervals described in Sect. 3.5.2. Thus, we assume that $d \sim \mathscr{U}[10000, 50000]$, $c_1 \sim \mathscr{U}[0,3]$, $e_1 \sim \mathscr{U}[1,6]$, $c_2 \sim \mathscr{U}[2,8]$ and $e_2 \sim \mathscr{U}[10,14]$. Then, we may use Algorithm 1 to estimate the required probability, where $\psi_A^k(d_1,x)$ designates the expected utility that the Attacker reaches, when the Defender implements d_1 and he implements attack x and the sampled parameters are $d^k, c_1^k, c_2^k, e_1^k, e_2^k$.

In our particular case, with $N = 10{,}000$, we obtain $\hat{p}(a|d_1) = 0.406$ (and, consequently, $\hat{p}(\bar{a}|d_1) = 0.594$). Similarly, $\hat{p}(a|d_2) = 0.764$ and $\hat{p}(\bar{a}|d_2) = 0.236$. Then, we have $\psi(d_1) = -14.7$, $\psi(d_2) = -81.2$ and $\psi(d_3) = -20.08$ and the optimal ARA defense d^*_{ARA} is d_1, which is different to d^*_{GT}.

Algorithm 1: Estimating $p(a|d_1)$

```
p = 0;
for k ← 1 to N do
    Sample d^k, c_1^k, c_2^k, e_1^k, e_2^k;
    if ψ_A^k(d_1, a) ≥ ψ_A^k(d_1, ā) then
        p = p + 1;
p̂(a|d_1) = p/N;
```

3.5.4 Robustness of the ARA Solution

We consider now the robustness of the ARA solution. For that, we consider classes of beta distributions with the same support than the corresponding parameters. As an example, for d, we shall assume that $d \sim \beta[o_1, o_2]$ over the interval $[10000, 50000]$, with $o_1 \in [0.5, 1.5]$, $o_2 \in [0.5, 1.5]$. Similarly, for the other parameters we use beta distributions over the previous intervals, with parameters as in Table 3.3, where the first parameter of the beta distribution is uniform over [LL, LU] and the second parameter of the beta distribution is uniform over [UL, UU].

We sample 100 times from such distributions and repeat the procedure in Sect. 3.5.3. Then, the estimated probabilities of each defense being optimal, in the ARA sense, would be, respectively, $\hat{p}(d_1) = 1$, $\hat{p}(d_2) = 0$ and $\hat{p}(d_3) = 0$. Therefore, d_1 seems clearly the most likely alternative for being optimal.

The regrets when we implement various solutions, are respectively, 0 for d_1, 37.91 for d_2 and 8.54 for d_3. Thus, the minimum regret defense is d_1.

Table 3.3: Upper and lower limits for the parameters of the involved beta distributions

Parameter	LL	LU	UL	UU
c_1	0.5	1.5	0.5	1.5
c_2	0.5	1.5	0.5	1.5
e_1	0.5	1.5	0.5	1.5
e_2	0.5	1.5	0.5	1.5

3.6 Discussion

Adversarial Risk Analysis is an emergent paradigm when supporting a decision maker who faces adversaries and such that the consequences are random and depend on the actions of all participating agents. The prevalent paradigm in this area is Game Theory. In this chapter, we have provided a framework for robustness analysis in this area.

The approach we have followed is:

- Under common knowledge assumptions compute the game theoretic solution. Perform a robust analysis for such solution. If it is stable, such solution may be used with confidence and we do not require further analysis.
- Otherwise, perform an ARA. Undertake a robust analysis for the ARA solution. If it is stable, the ARA solution may be used with confidence and the analysis stops. Otherwise, gather more data and/or refine the relevant classes, eventually declaring the robustness of the ARA solution. If not sufficient, move towards next stage.
- Undertake a minimum regret (or other robust) concept.

We have illustrated it with two simple models, the sequential defend-attack and the simultaneous defend-attack, but the ideas would extend to more complex ARA models. Similarly, we have assumed that the attacker was maximising expected utility but the ideas may be translated to other attacker rationalities, as in [21].

There are many other sensitivity analysis questions relevant in ARA. For example, we mentioned above the recursive assessment required in the simultaneous game, which may be expressed as follows, see [17]:

Algorithm 2: Recursive assessment required in the simultaneous game

for $i \leftarrow 1$ **to** ∞ **do**

Find $\Pi_{D^{i-1}}(A^i)$ by solving

$$A^i \mid D^i \sim \arg\max_{a \in \mathscr{A}} \sum_{d \in \mathscr{D}} \left[\sum_{s \in \{0,1\}} U_A^i(a,s) \;\; P_A^i(S = s \mid d, a) \right] \Pi_{A^i}(D^i = d)$$
$$\text{with } (U_A^i, P_A^i) \sim F^i$$

Find $\Pi_{A^i}(D^i)$ by solving

$$D^i \mid A^{i+1} \sim \arg\max_{d \in \mathscr{D}} \sum_{a \in \mathscr{A}} \left[\sum_{s \in \{0,1\}} U_D^i(d,s) \;\; P_D^i(S = s \mid d, a) \right] \Pi_{D^i}(A^{i+1} = a)$$
$$\text{with } (U_D^i, P_D^i) \sim G^i$$

$i = i + 1$;

This hierarchy would stop when the Defender lacks the information necessary to assess the distribution F^i or G^i associated with the decision analysis of A^i and D^i, respectively. At this point, the Defender would assign an unconditional probability distribution over A^i or D^i, respectively, without going deeper into the hierarchy, summarizing all the information she might have through the direct assessment of $\Pi_{D^{i-1}}(A^i)$ or $\Pi_{A^i}(D^i)$, as might correspond. Should she have no additional information to do so, she could assign a noninformative distribution, see [3].

However, climbing up one level in the hierarchy entails a lot of effort. We could question whether this is worth it by using value of information types of computation.

Acknowledgements

DRI, CA and JG are grateful to the MINECO MTM2014-56949-C3-1-R and the AXA-ICMAT Chair in Adversarial Risk Analysis. The authors are grateful to the support of the COST IS1304 Action on Expert Judgement.

References

1. Banks, D.L., Aliaga, J.M.R., Ríos Insua, D.: Adversarial Risk Analysis, vol. 343. CRC Press, Boca Raton, FL (2015)
2. Berger, J.O., Ríos Insua, D., Ruggeri, F.: Bayesian robustness. In: Ríos Insua, D., Ruggeri, F. (eds.) Robust Bayesian Analysis, pp. 1–32. Springer, New York (2000)
3. French, S., Ríos Insua, D.: Statistical Decision Theory. Wiley, New York (2000)
4. Gibbons, R.: A Primer in Game Theory. Harvester Wheatsheaf, Hemel Hempstead (1992)
5. Hargreaves-Heap, S., Varoufakis, Y.: Game Theory: A Critical Introduction. Routledge, Abingdon (2004)
6. Harsanyi, J.C.: Games with incomplete information played by "bayesian" players, part i. The basic model. Manag. Sci. **14**(3), 159–182 (1967)
7. Harsanyi, J.C.: Games with incomplete information played by "bayesian" players, part ii. Bayesian equilibrium points. Manag. Sci. **14**(5), 320–334 (1968)
8. Harsanyi, J.C.: Games with incomplete information played by "bayesian" players, part iii. The basic probability distribution of the game. Manag. Sci. **14**(7), 486–502 (1968)
9. Kardes, E.: Robust Stochastic Games and Applications to Counter-Terrorism Strategies. Center for Risk and Economic Analysis of Terrorism Events, University of Southern California, Los Angeles (2005)
10. Lippman, S.A., McCardle, K.F.: Embedded Nash bargaining: risk aversion and impatience. Decis. Anal. **9**(1), 31–40 (2012)
11. McLay, L., Rothschild, C., Guikema, S.: Robust adversarial risk analysis: a level-k approach. Decis. Anal. **9**(1), 41–54 (2012)
12. Merrick, J., Parnell, G.S.: A comparative analysis of PRA and intelligent adversary methods for counterterrorism risk management. Risk Anal. **31**(9), 1488–1510 (2011)
13. Nau, R.F.: Joint coherence in games of incomplete information. Manag. Sci. **38**(3), 374–387 (1992)
14. O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T.: Uncertain Judgements: Eliciting Experts' Probabilities. Wiley, Hoboken, NJ (2006)
15. Parnell, G., Banks, D., Borio, L., Brown, G., Cox, L., Gannon, J., Harvill, E., Kunreuther, H., Morse, S., Pappaioanou, M., et al.: Report on methodological improvements to the department of homeland security's biological agent risk analysis National Academies Press, Washington DC (2008)
16. Raiffa, H., Richardson, J., Metcalfe, D.: Negotiation Analysis: The Science and Art of Collaborative Decision Making. Harvard University Press, Cambridge, MA (2002)
17. Rios, J., Ríos Insua, D.: Adversarial risk analysis for counterterrorism modeling. Risk Anal. **32**(5), 894–915 (2012)
18. Ríos Insua, D., Ruggeri, F.: Robust Bayesian Analysis. Springer, New York (2000)

19. Ríos Insua, D., Ruggeri, F., Vidakovic, B.: Some results on posterior regret γ-minimax estimation. Stat. Decis. **13**, 315–331 (1995)
20. Ríos Insua, D., Rios, J., Banks, D.: Adversarial risk analysis. J. Am. Stat. Assoc. **104**(486), 841–854 (2009)
21. Ríos Insua, D., Banks, D., Rios, J.: Modeling opponents in adversarial risk analysis. Risk Anal. (2015)
22. Rothschild, C., McLay, L., Guikema, S.: Adversarial risk analysis with incomplete information: a level-k approach. Risk Anal. **32**(7), 1219–1231 (2012)
23. Sevillano, J.C., Rios Insua, D., Rios, J.: Adversarial risk analysis: the Somali pirates case. Decis. Anal. **9**(2), 86–95 (2012)
24. Stahl, D.O., Wilson, P.W.: On players models of other players: theory and experimental evidence. Games Econ. Behav. **10**(1), 218–254 (1995)
25. Von Winterfeldt, D., O'Sullivan, T.M.: Should we protect commercial airplanes against surface-to-air missile attacks by terrorists? Decis. Anal. **3**(2), 63–75 (2006)
26. Wang, C., Bier, V.M.: Expert elicitation of adversary preferences using ordinal judgments. Oper. Res. **61**(2), 372–385 (2013)
27. Zhuang, J., Bier, V.M.: Balancing terrorism and natural disasters-defensive strategy with endogenous attacker effort. Oper. Res. **55**(5), 976–991 (2007)

Chapter 4
From Statistical Decision Theory to Robust Optimization: A Maximin Perspective on Robust Decision-Making

Moshe Sniedovich

Abstract As attested by the prevalence of worst-case-based robustness analysis in many fields, Wald's maximin paradigm (circa 1940) plays a central role in the broad area of decision-making under uncertainty. The objective of this chapter is therefore twofold. First, to examine the basic conceptual and modeling aspects of this ostensibly intuitive, yet controversial paradigm, so as to clarify some of the issues involved in its deployment in decision-making in the face of a non-probabilistic uncertainty. Second, to elucidate the differences between this paradigm and other maximin paradigms, such as those used in *error analysis* and *game theory*.

We thereby chart the journey of this paradigm from the field of *statistical decision theory* to that of modern *robust optimization*, highlighting its use in the latter, as a tool for dealing with both *local* and *global* robustness. We also look briefly at the relationship between probabilistic and worst-case-based robustness analysis.

4.1 Introduction

No more than a cursory examination of the literature on the employment of the apparently immediately intelligible concept *robustness* in such diverse fields as statistics, decision theory, control theory, optimization, economics, engineering, machine learning, and so on, suffices to see that the decision-making models used in these fields are often *minimax*, or *maximin* models.

M. Sniedovich (✉)
Department of Mathematics and Statistics, University of Melbourne, Melbourne, VIC 3010, Australia
e-mail: moshe@unimelb.edu.au

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_4

This is hardly surprising, considering that the robustness criterion that is often invoked in these and other fields is based on a *worst-case* approach to *uncertainty* and *variability,* and that maximin models provide the most obvious formal framework for implementing such an approach in a decision-making environment.

That said, it ought to be emphasized straightaway that maximin models are used for various other purposes as well. This means that it is important to be clear in what way, if any, do maximin models used in worst-case-based robustness analysis differ from maximin models that are used for other ends.

For this reason, we focus in this discussion not only on the mathematical structure of generic maximin models, but also on their conceptual dimension, that is, on the *conceptual* purport of the abstract mathematical objects comprising these models and on the manner in which they hang together.

Our discussion is thus organized along the following lines. We begin with a brief informal description of the *decision problem* that is of concern to us in this discussion, and we explain how the maximin paradigm approaches such problems:

Section 4.2: The fundamental decision problem
Section 4.3: Wald's maximin paradigm
Section 4.4: Maximin models at a glance
Section 4.5: The Wald factor

Next, we outline the kinds of robustness that are furnished by Wald-type maximin models and we illustrate some of the issues associated with the use of such models:

Section 4.6: Robustness
Section 4.7: A perspective on robust decision-making

We conclude the discussion with some general comments on Wald's maximin paradigm, notably its role and place in robust-decision-making:

Section 4.8: Can Wald's maximin paradigm save the world?

And so, our first task in charting the maximin paradigm's journey from *statistical decision theory* (circa 1940) to modern *robust optimization* is to examine the fundamental decision problem that is at the heart of the discussion on this issue.

4.2 The Fundamental Decision Problem

Under consideration in this discussion are situations where the objective is to identify a *decision* $d \in D$ that is *robust* against variations in the *state* of the system, $s \in S$. By this we mean a *decision* $d \in D$ that performs "well" (relatively to other decisions) with respect to given *performance criteria*, in the face of variations in the value of $s \in S$. We refer to set D as the *decision space* and to set S as the *state space*. We further assume that the decision maker has full control over the value of the decision variable d, but no control over the value of the state variable, s.

Of particular interest to us here are situations where the state variable represents *uncertainty,* in which case we assume that the value of s is selected from S by *Nature* in response to the decision maker selecting a decision d from D. One of the implications of this is that the decision maker is ignorant as to which value of s will be realized in response to his/her selection of a decision d from D.

To simplify matters, let us focus on two performance criteria, namely *payoff* and *feasibility*. In this framework the *payoff* is assumed to be a numeric scalar such that the larger it is the better; and *feasibility* amounts to satisfying *constraints* imposed on the (decision, state) pairs. Satisfying these constraints is assumed to be preferable to violating them. Thus, a robust decision is one that (relative to other decisions) yields a large payoff and satisfies the constraints under consideration over a large subset of the state space S.

In this framework then, the fundamental decision problem faced by the decision maker can be stated informally as follows:

- Select a decision $d \in D$ that is robust against variations in the value of s over S with respect to the payoffs and constraints under consideration such that (relative to other decisions) it yields a large payoff and satisfies the constraints over a large subset of S .

To render the notation and terminology simple, let $O(d,s)$ denote the *outcome* generated by decision d and state s, so in the above simple framework $O(d,s) = (p,c)$ where p denotes the payoff generated by d and s and c is an indicator stipulating whether the pair (d,s) satisfies the performance constraints. Regarding the latter, let $\vee$ indicate that the constraints are *satisfied* and let $\times$ indicate that at least one of the constraints is *violated*. To illustrate, consider the situation where d can take three values, s can take five values, and the respective outcomes are as follows:

$O(d,s)$	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$	$s^{(4)}$	$s^{(5)}$
$d^{(1)}$	$(55,\times)$	$(43,\vee)$	$(18,\vee)$	$(63,\vee)$	$(37,\times)$
$d^{(2)}$	$(38,\vee)$	$(22,\vee)$	$(11,\vee)$	$(12,\vee)$	$(10,\vee)$
$d^{(3)}$	$(85,\vee)$	$(83,\times)$	$(23,\times)$	$(72,\vee)$	$(50,\times)$

(4.1)

By inspection, decision $d^{(2)}$ performs very well with respect to the constraints and decision $d^{(3)}$ performs very well with respect to the payoff. The questions naturally arising are therefore these:

- How do we measure robustness in this case?
- How do we determine the best (optimal) decision in this case?

Wald's maximin paradigm offers a simple (some might argue simplistic) recipe to deal with questions of this type: *assume the worst!*

4.3 Wald's Maximin Paradigm

Wald's maximin paradigm exemplifies an approach to dealing with uncertainty and variability that, however controversial, can arguably be described as natural and intuitive: *in the face of uncertainty . . . assume the worst!*

Thus, putting this approach into action, the paradigm prescribes ranking decisions according to their worst performance, whereupon the decision that yields the best (over decisions) worst (over states) outcome, is deemed the best (optimal) decision. In other words, the *decision rule* articulated by this paradigm can be stated in plain language as follows [34, pp. 152–153]:

> The maximin rule tells us to rank alternatives by their worst possible outcomes: we are to adopt the alternative the worst outcome of which is superior to the worst outcome of the others.

To illustrate its working, consider how the decision problem whose outcomes are specified by (4.1) would be handled on grounds of this rule. Central to its implementation is the supposition of a tradeoff between *payoffs* and *feasibility,* where the underlying idea is, as in *optimization theory*, that feasibility (constraints satisfaction) takes precedence over payoff. This means that an outcome $(p,\vee)$ is superior to an outcome $(p',\times)$, regardless of the values of the payoffs p and p'. And in this vein, the worst outcomes (WO) pertaining to the three decisions considered in (4.1) are as indicated by the last column of the following table:

$O(d,s)$	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$	$s^{(4)}$	$s^{(5)}$	WO
$d^{(1)}$	$(55,\times)$	$(43,\vee)$	$(18,\vee)$	$(63,\vee)$	$(37,\times)$	$(37,\times)$
$d^{(2)}$	$(38,\vee)$	$(22,\vee)$	$(11,\vee)$	$(12,\vee)$	$(10,\vee)$	$(10,\vee)$
$d^{(3)}$	$(85,\vee)$	$(83,\times)$	$(23,\times)$	$(72,\vee)$	$(50,\times)$	$(23,\times)$

(4.2)

The implication is that, according to the precepts of Wald's maximin paradigm, the best decision (alternative) is $d^{(2)}$, the second-best decision is $d^{(1)}$ and the third-best decision is $d^{(3)}$. The respective worst outcomes are $(10,\vee)$, $(37,\times)$ and $(23,\times)$.

While "outcome tables", such as those displayed above, give an immediate illustration of the basic ideas informing Wald's maximin paradigm and the decision rule it proposes, the full picture of this paradigm's prowess is brought to light by its operation in the framework of *constrained optimization* problems. So, to set the scene for a discussion of this point, we formulate a generic maximin model, and two of its instances, that are particularly suitable for this purpose.

As a prelude, we note that for simplicity of exposition, we assume that the optimization problems under consideration are (unless explicitly stated otherwise) well-behaved in the sense that they possess (global) optimal solutions. With this in mind, let us examine three generic maximin models.

4.4 Maximin Models at a Glance

Let $f(d,s)$ denote the *payoff* generated by decision $d \in D$ and state $s \in S$ and let *constraints*(d,s) denote the list of *constraints* imposed on (d,s) pairs. Formally, we can regard f as a real-valued function on $D \times S$. Since in many applications only a subset of S is relevant with regard to a particular decision, it is instructive to consider the case where each decision has its own set of states.

Hence, let $S(d)$ denote the set of all the possible/plausible values of s that can be generated by decision d and let $S := \cup_{d \in D} S(d)$. Then, as feasibility (constraints satisfaction) has priority over payoffs, an implementation of the maximin decision rule would yield the following maximin model:

$$z^* := \max_{d \in D} \min_{s \in S(d)} \{f(d,s) : \textit{constraints}(d,s), \forall s \in S(d)\}. \tag{4.3}$$

Note that in this model the $\forall s \in S(d)$ clause is a *worst-case requirement* which gives expression to the priority of feasibility over payoffs. Thus, granted this priority, a decision $d' \in D$ that satisfies the constraints for all $s \in S(d')$ is strictly superior to any decision $d \in D$ that violates (at least one of) these constraints for some $s \in S(d)$, regardless of the payoffs generated by d' and d.

I call maximin models of this type *Full Monty* maximin models for the obvious reason that they seek robustness with respect to both payoff and constraints. The former is represented by the $\min_{s \in S(d)}$ operation and the latter by the $\forall s \in S(d)$ requirement.

Now, consider the instance of this model that is characterized by the property that the payoff $f(d,s)$ is independent of s. In this case, the payoff generated by decision d is denoted $f(d)$ rather than $f(d,s)$. Note that in this case the operation $\min_{s \in S(d)}$ is superfluous in (4.3). Hence, this instance of (4.3) takes this form:

$$z^* := \max_{d \in D} \{f(d) : \textit{constraints}(d,s), \forall s \in S(d)\}. \tag{4.4}$$

I call maximin models of this type *state-free-payoff* maximin models to underscore the fact that in these models payoffs are independent of the state variable, and therefore robustness against variations in the value of s is sought only with respect to the constraints.

And, to complete this sketch, consider what to many readers probably epitomizes the maximin paradigm, namely the generic maximin model that is devoid of explicit constraints on (d,s) pairs. This is the instance of (4.3) where the list *constraints*(d,s) is empty. This model has this simple form:

$$z^* := \max_{d \in D} \min_{s \in S(d)} f(d,s). \tag{4.5}$$

I call maximin models of this type *textbook* maximin models. They represents situations where robustness is sought only with respect to payoffs.

Having said all that, it should be emphasized that, from a modeling point of view, these three generic maximin models are *interchangeable.* That is, each one of these models can be reformulated so as to assume the format of anyone of the other two. For instance, the generic *Full Monty* maximin model can be rewritten as follows:

$$z^* := \max_{d\in D, t\in\mathbb{R}} \{t : t \le f(d,s), constraints(d,s), \forall s \in S(d)\} \tag{4.6}$$

where $\mathbb{R}$ denotes the real-line. Clearly, this is a *state-free-payoff* maximin model.

4.4.1 Security Levels

In the idiom of *decision theory,* the worst payoff pertaining to decision $d \in D$, yielded by the inner minimization problem in a maximin model, is termed *security level.* Thus, the *security level* of decision d is equal to the smallest payoff pertaining to this decision, assuming that the decision satisfies the worst-case constraints, if any. Hence, in the context of *textbook* maximin models the *security level* of decision d is defined as follows:

$$SL(d) := \min_{s\in S(d)} f(d,s)\ ,\ d \in D. \tag{4.7}$$

In cases where there are explicit constraints on (d,s) pairs, let $D^{(a)}$ denote the subset of D whose elements satisfy the constraints for all the states pertaining to the respective decision, namely define

$$D^{(a)} := \{d \in D : constraints(d,s), \forall s \in S(d)\}. \tag{4.8}$$

Observe that if there are no explicit constraints on the (d,s) pairs, as in the case of the *textbook model*, then $D^{(a)} = D$. Hence, in the context of *Full Monty* maximin models, the security level of decision $d \in D$ is defined as follows:

$$SL(d) := \begin{cases} \min\limits_{s\in S(d)} f(d,s) & ,\ d \in D^{(a)} \\ \text{inadmissible} & ,\ d \notin D^{(a)} \end{cases},\ d \in D. \tag{4.9}$$

Needless to say, if all the decisions are inadmissible, namely if $D^{(a)}$ is empty, then the maximin problem has no feasible, let alone optimal, solution. Also, observe that if $D^{(a)} = D$ then the constraints are superfluous and therefore can be ignored. For this reason we focus on cases where $D^{(a)}$ is a non-empty proper subset of D.

Note that in the context of *state-free-payoff* models, the expression for $SL(d)$ simplifies to

$$SL(d) := \begin{cases} f(d) & ,\ d \in D^{(a)} \\ \text{inadmissible} & ,\ d \notin D^{(a)}. \end{cases} \tag{4.10}$$

In all these cases the maximin models can be written as follows:

$$z^* := \max_{d \in D} SL(d). \tag{4.11}$$

In this setting, the operation $\max_{d \in D}$ ignores inadmissible decisions. So, if all the decisions are inadmissible, the problem has no optimal solution. Hence, the formulation of the maximin rule can be fine-tuned as follows:

Maximin rule:
Rank admissible decisions according to their *security level:* the larger the security level the better. Hence, select an admissible decision whose security level is the largest.

4.4.2 Optimal Solutions

To be clear on what constitutes an optimal solution to a maximin problem of the type examined in this chapter, observe that such a solution is a pair (d^*, s^*) such that $d^* \in D^{(a)}$, $s^* \in S(d^*)$ and

$$f(d^*, s^*) = SL(d^*) = \max_{d \in D} SL(d). \tag{4.12}$$

Note also that in the case of state-free-payoff maximin models, all the elements of $S(d^*)$ are optimal with respect to d^*.

And as a final note, to avoid becoming bogged down by technical issues related to the existence of optimal solutions for maximin problems, this discussion is predicated on the supposition that the maximin problems under consideration have optimal solutions.

4.4.3 A Constrained Optimization Perspective

As we saw above, *Full Monty* and *textbook* maximin models can be written as *state-free-payoff* maximin models. The question therefore arising is: what renders maximin problems different from "conventional" constrained optimization problems, namely problems seeking the optimization of a real-valued *objective function* subject to a *finite* list of *constraints* on the decision variable. To address this issue, let us juxtapose the following two generic optimization problems against each other:

Maximization problem	State-free-payoff maximin problem
$\max_{d \in D} \{f(d) : constraints(d)\}$	$\max_{d \in D} \{f(d) : constraints(d, s), \forall s \in S(d)\}$

(4.13)

where $constraints(d)$ is a list of constraints imposed on the decision variable d (in addition to the domain constraint $d \in D$).

Clearly, the two problems are similar in that both require the maximization of a real-valued objective function subject to constraints imposed on the decision variable. The difference between them is in the structures of their respective constraints. Thus, while the constraints of the maximization problem apply only to the decision variable d, the constraints of the state-free-payoff maximin model apply to (d,s) pairs, what is more, they are required to be satisfied *for all* $s \in S(d)$.

The upshot of this is that in cases where the sets $S(d), d \in D$ consist of *infinitely many* elements, the *state-free-payoff* maximin problem has *infinitely many* constraints. Namely, the maximin problem is a *semi-infinite* optimization problem [20, 43]. Furthermore, given the possible *interaction* between d and s, keep in mind that constraints that are say *linear* with respect to d and *linear* with respect to s, may not necessarily be linear with respect to (d,s). This means that incorporating a state variable in a system of linear constraints imposed on the decision variable d may render the constraints nonlinear (with respect to (d,s) pairs) in the context of the maximin problem. Similarly, incorporating a state variable in a system of say *convex* constraints on the decision variable d, may render the constraints *non-convex* (with respect to (d,s) pairs) in the context of the maximin problem.

For these reasons, the state-free-payoff maximin counterpart of a conventional optimization problem is often far more difficult to solve than the conventional optimization problem.

4.5 The Wald Factor

The fact that the maximin models considered in this chapter are discussed under the rubric *Wald's maximin paradigm* does not in any way imply that the maximin concept, principle, or paradigm, originated in the work of the mathematician Abraham Wald (1902–1951). Clearly, this is not the case, even if the proliferation, since the 1940s, in the use of maximin and minimax models in applied mathematics and other fields, might give the impression that *maximin,* as a concept, is a recent invention. The fact of the matter is that its use can be traced back at least to *error analysis*, namely, at least to Leonhard Euler (1707–1783). Thus, for an idea of the distant roots of these models, consider for instance the article *Origin of the theory of errors* by Sheynin [38] who argues that Johann Heinrich Lambert (1728–1777) "... should be given precedence over Gauss as the originator of the theory of errors ..." because, as [38, p. 1004] points out:

> (f) An enunciation of the "minimax" principle (minimization of the maximum residual error in geodetic adjustments—the minimum being sought among all possible solutions .... But Lambert confessed that he did not know how to use this principle "in a general manner, without many devious ways" (*auf eine allgemeine Art, und ohne viele Umwege*). The use of this principle in a rudimentary form for solving a redundant system of linear algebraic equations should be credited to Euler....

In a nutshell, the *maximin paradigm* under discussion here is the one that was first introduced by the mathematician Abraham Wald (1902–1950) into *statistical decision theory* in 1939, from whence it entered *decision theory* in the 1950s, and subsequently many other disciplines, including *robust decision-making* and *robust optimization*.

To be precise, it is the paradigm that was introduced in the seminal article *Contributions to the theory of statistical estimation and testing hypothesis* where [46, p. 305] argued as follows (emphasis added):

> There exist in general many systems M_s which are admissible relative to the weight function given. The question arises as to how can we distinguish among them. Denote by r_{M_s} the **maximum** of the risk function corresponding to the system M_s of regions and to the given weight function. If we do not take into consideration a priori probabilities of θ, then **it seems reasonable** to choose that system M_s for which r_{M_s} becomes **minimum.**

It therefore must not be conflated with the classical maximin paradigm of zero-sum 2-person games which became popular with the publication of [45] seminal book *Theory of Games and Economic Behavior*.

The importance of Wald's paradigm was immediately recognized, as indeed attested, for example, by Morgenstern [28, pp. 355–356]:

> In practical and scientific affairs there is always need to decide upon courses of action, as a rule on the basis of incomplete information. This problem is one of whether to accept or to reject a particular course, or, more generally, to find the optimum course from a wider set of possibilities, where a wrong choice results in a loss to be suffered. The best that can be hoped for is to minimize the maximum loss. This principle of action is known as the minimax principle. Wald introduced it into statistics and has given it basic importance for the theory of statistical decision functions. It has been claimed that "it is the only rule of comparable generality proposed since [that of] Bayes' was published in 1763."[3] The minimax principle is central in the theory of games, and, when that theory was published in its present form in 1944, it attracted Wald immediately. In various papers he set forth the applicability of the theory of two-person zero-sum games to the theory of statistics and extended certain game theoretical results, notably generalizing the main theorem to the case of a denumerably infinite set of strategies.

[3] Savage [36, p. 59]

For the record, though, it is important to note that no reference whatsoever is made in [46] to *game theory*, and the idea of minimizing the maximum risk advanced in this article is not based on any game theoretic concept, or argument. Only 6 years later, in the article *Statistical decision functions which minimize the maximum risk*, did [47] discuss the connection between the model proposed in [46] and zero-sum, 2-person games.

The central difference between Wald-type maximin models and maximin models associated with zero-sum, 2-person games is that the former do not postulate any *equilibrium* conditions. Therefore, a solution to a Wald-type maximin problem is not required to be a *saddle point*. To illustrate, consider the following simple case consisting of 3 decisions, 5 states, and the payoffs are as follows:

$f(d,s)$	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$	$s^{(4)}$	$s^{(5)}$	SL
$d^{(1)}$	55	43	48	63	37	37
$d^{(2)}$	88	92	71	10	20	10
$d^{(3)}$	85	83	47	72	50	47

(4.14)

The SL column contains the *security levels* pertaining to the three decisions where, as we saw above, in the maximin context, the assumption is that the larger the payoff the better.

Thus, from a Wald's maximin paradigm standpoint, the optimal solution to the maximin problem defined by this payoff table is $(d^*, s^*) = (d^{(3)}, s^{(3)})$, yielding a payoff of 47. But, from a zero-sum, 2-person game standpoint, this pair is not in equilibrium, that is $(d^{(3)}, s^{(3)})$ is not a *saddle point:* despite it being the smallest element in its row, it is not the largest element in its column. Hence, $(d^{(3)}, s^{(3)})$ does not constitute a solution to the zero-sum, 2-person maximin game defined by this payoff table. But more than that, this game has no *saddle point,* it therefore has no optimal solution that is a *pure strategy.*

Conceptually, Wald-type maximin and minimax models can be regarded as games between two players: the decision maker and Nature which embodies the decision maker's opponent. Such a game consists of the following moves:

Step 1. The decision maker moves first by selecting a decision $d \in D$.
Step 2. In response, Nature selects a state $s \in S(d)$.
Step 3. Then, an outcome $O(d,s)$ is realized.

That is, in this game the decision maker, who "plays" first, seeks the *best* outcome, whereas Nature who "plays" second, seeks the *worst* outcome. So if the decision maker selects say decision $d' \in D$, then Nature will select a state $s' \in S(d')$ that yields the worst outcome $O(d', s')$ over all $s \in S(d')$. In symbols,

$$z^* := \underset{d \in D}{\text{best}} \ \underset{s \in S(d)}{\text{worst}} \ O(d,s) \tag{4.15}$$

where the operation $\text{best}_{d \in D}$ determines the best outcome over the set of decisions available to the decision maker, and the operation $\text{worst}_{s \in S(d)}$ determines the worst outcome over the set of states associated with decision d.

In cases where the outcomes are expressed in terms of *payoffs* and *constraints*, as done in Sect. 4.2, using the term *pessimization* as the antonym of *optimization* (see [4, 7, 12, 24, 30, 31]), enables formulating the following abstract Wald-type decision-making model:

$$z^* := \underset{d \in D}{\text{opt}} \ \underset{s \in S(d)}{\text{pes}} \ \{f(d,s) : \textit{constraints}(d,s), \forall s \in S(d)\} \tag{4.16}$$

where "pes" denotes the converse of "opt", namely

$$\text{opt} = \min \iff \text{pes} = \max \tag{4.17}$$

$$\text{opt} = \max \iff \text{pes} = \min . \tag{4.18}$$

Observe that the models specified by (4.15) and (4.16) encompass both minimax (opt=min) and maximin (opt=max) models.

In view of what we have seen thus far, we can now sum up the essential features that distinguish Wald-type maximin/minimax models from other minimax/maximin models. In this discussion then the term *Wald-type* designates maximin/minimax models where:

- The state variable s represents "non-probabilistic" uncertainty.
- The decision maker "plays" first.
- Nature's sole goal is to inflict the greatest possible harm on the decision maker.

The phrase "non-probabilistic" is designed to indicate that not only is the uncertainty in the true value of the state variable s not quantified probabilistically, no measure of "likelihood", or "chance", or "plausibility", or "belief" is associated with s. The assumption is that all that is known about the (unknown) true value of s is that it is an element of $S(d)$. Also note that the notation $s \in S(d)$ implies that the uncertainty in the true value of s may be affected by the decision $d \in D$ made by the decision maker in the sense that the set of possible/plausible values of this true value may depend on d.

As for Nature's hostile antagonistic stance towards the decision maker, interestingly the upshot of this is that it effectively eliminates the uncertainty as to the game's outcome. Namely, if the decision maker selects decision $d \in D$ then in the context of (4.16) there is no uncertainty whatsoever regarding the game's outcome: it is certain to be equal to

$$SL(d) := \underset{s \in S(d)}{\text{pes}} \{f(d,s) : constraints(d,s), \forall s \in S(d)\} \tag{4.19}$$

which is the *security level* of decision d.

The implication is that it is important to distinguish between the following two cases:

$$SL(d) = \begin{cases} \underset{s \in S(d)}{\text{pes}} \; f(d,s) & , \; constraints(d,s), \forall s \in S(d) \\ \text{inadmissible} & , \; \text{otherwise} \end{cases} , \; d \in D. \tag{4.20}$$

This means that the generic model specified by (4.16) can be written as follows:

$$z^* := \underset{d \in D}{\text{opt}} \; SL(d) \tag{4.21}$$

assuming that here the operation $\text{opt}_{d \in D}$ ignores inadmissible decisions.

And to sum it all up, Wald-type models are maximin and minimax models where the state variable, s, represents *uncertainty*, namely it is a parameter of the decision-making model whose true value is uncertain.

4.6 Robustness

Before we proceed to examine the kind of robustness that Wald-type maximin models are concerned with, let us first examine the meaning of the concept "robustness" by considering, albeit very briefly, how this concept functions in a number of disciples.

A quick scan of publications from various fields suggest that although the concept "robustness" may be given specific definitions in different disciplines to reflect their specific concerns and objectives, in essence the meanings that the various definitions seek to convey build on the purport of the concept "robustness" that is familiar to us from ordinary language. Thus, definitions of "robustness" from various disciplines tend to explain "robustness" in terms of "an ability to withstand ...", "resilience to ...", "insensitivity to ...", and so on. And to illustrate, consider the definitions articulated in the following publications.

According to [25, p. 126], the Eurocode (EN 1991-7-1 Clause 1.5.14) defines *robustness of structures* as follows:

> Robustness is the ability of a structure to withstand events like fire, explosions, impact or the consequences of human error, without being damaged to an extent disproportionate to the original cause.

And according to [33, p. 183]:

> Robustness as a mathematical concept was introduced in a short paper by Andronov and Pontriagin[1] and developed in a later book by Andronov et al.[2] The original definition was of the local type, i.e. it characterized variations in the behaviour of a dynamic system under small variations in its velocity field. The robustness was interpreted both as a qualitative, geometrical property (invariance of a topological structure) and as a qualitative, analytical property (uniform continuity of solutions depending on a parameter characterizing the small variations).

[1] = Andronov and Pontriagin [1]; [2] = Andronov et al. [2]

And the synopsis of the article *Permutation theory in the derivation of robust criteria and the study of departures from assumption* [9, p. 1] asserts the following:

> In the practical circumstances in which statistical procedures are applied, little is usually known of the validity of assumptions such as normality of the error distribution. Procedures are required which are "robust" (insensitive to changes in extraneous factors not under test) as well as powerful (sensitive to specific factors under test). Permutation theory, which provides one method for deriving robust criteria, is discussed and applied to the problem of comparing variances.

We call attention to the fact that although our main concern in this chapter is with *worst-case-based* robustness, it is important to keep in mind that the latter is rivaled by *probabilistic robustness* as indicated in the article *Probabilistic robustness analysis—risks, complexity, and algorithms* [11, p. 2693] where we read the following:

> In recent years, a number of researchers have proposed probabilistic control methods for overcoming the computational complexity and conservatism of the deterministic worst-case robust control framework ...
>
> The philosophy of probabilistic control theory is to sacrifice cases of extreme uncertainty. Such a paradigm has led to the concept of confidence degradation function (originated by Barmish, Lagoa, and Tempo [2]), which has been demonstrated to be extremely powerful for the robustness analysis of uncertain systems.

[2] = Barmish et al. [3]

And with this in mind, let us now examine the robustness that is sought by Wald-type maximin models.

4.6.1 Worst-Case-Based Robustness

As we noted already, Wald's worst-case-based maximin paradigm exemplifies a rather grim, indeed pessimistic view of uncertainty, which however controversial, can arguably be described as natural and intuitive. Indeed, one might further argue that such an approach underlies time old adages such as *when in doubt assume the worst!* and *hope for the best but plan for the worst!* And that it is most certainly reflected in the following stanza from William Shakespeare's *Julius Caesar* (Act 5 Scene 1) which serves as an epigraph to Rustem and Howe's [35] book *Algorithms for Worst-case Design and Applications to Risk Management* that is dedicated to "those who have suffered the worst case":

The gods to-day stand friendly, that we may,
Lovers in peace, lead on our days to age!
But, since the affairs of men rest still incertain
Let's reason with the worst that may befall.

To be sure, one might counter that, for all our claimed inclination to "reason with the worst that may befall", experience with natural and man-made disasters shows time and again, that individuals, organizations, societies and nations are very often ill-prepared for events that are far less extreme than worst-case scenarios such as say the so-called 100-year flood—let alone Noah's flood!

This, in large part, is due to the fact that one of the most vexing difficulties posed by a worst-case-based approach to uncertainty is that its implementation can come at an exorbitant cost and may demand radical changes in our way of life.

We discuss these and related issues associated with the implementation of worst-case-based robustness models in the ensuing sections. Prior to that, let us clarify in broad terms how Wald's maximin paradigm, as an exponent, indeed the *ultimate* exponent, of a worst-case-based approach to uncertainty, quantifies the intuitive concept "robustness".

Observe then that from the standpoint of Wald's maximin paradigm, the robustness of decision d is a measure of how well or how poorly does it perform if the worst state in $S(d)$ is realized. This worst-case performance is quantified by the

security level of decision d, namely by $SL(d)$, defined in Sect. 4.4.1. In other words, in the framework of Wald-type maximin models, the robustness of decision d against the uncertainty in the true value of s is equal to the *security level* of decision d. In this framework, the larger the *security level* the better, hence the optimal decision is that whose *security level* is the largest.

In parallel, in the framework of *minimax* models, the smaller the *security level* the better, hence the optimal decision is that whose *security level* is the smallest. Hence, it is only natural that, in the context of minimax models, the *security level* of decision d would be regarded as a measure of its *fragility* rather than its *robustness*. Minimax models thus rank decisions according to their *fragility*: the smaller the fragility the better.

In the next subsection we take up the meaning and practical aspects of the term *worst* in the framework of worst-case analysis including Wald-type maximin models.

4.6.2 How Bad Should Worst Be?

One of the precepts guiding the analysis of a decision's robustness against uncertainty in the true value of a parameter, is that the values of the parameter that are considered by the analysis for this purpose ought to give a sound representation of the parameter's *variability*. In the context of a worst-case analysis, this means of course that the values of the parameter that are examined by the analysis ought to give a proper representation of the "worst value" of the parameter.

But this, it would seem, is "easier said than done" because, as pointed out by Hart et al. [21, p. 18], the whole question is:

> How worst is "worst-case"? Is the scenario literally the very worst (possible or imaginable?) case. Or is it a "realistic" (how realistic?) worst case. Or does it represent some known centile (95^{th}, 99^{th} ?) of a distribution of cases?

To appreciate how this issue comes into play in the context of Wald-type maximin models, keep in mind that, as explained in Sect. 4.4, $S(d)$ denotes the set of all the possible/plausible values of s that can be generated by decision d. Therefore, the question as to how "realistic" is the worst-case analysis conducted by Wald-type maximin models is intimately connected to the question of how "realistic" are the sets $S(d), d \in D$.

And this question, one need hardly point out, goes straight to the centrally important issue of *mathematical modeling* in general, and *maximin modeling* in particular. The point is that an answer to this question would in the first place depend on the analyst, namely on his/her insights, experience and skills in the *art of maximin modeling*, not the paradigm itself. Thus, in practice, it would all come down to the analyst ensuring that the sets $S(d), d \in D$ are quantified properly.

We address this issue in Sect. 4.7.2. In the next section we take up a related matter, which is the ability of Wald's maximin paradigm to deal with both *local* and *global* robustness.

4.6.3 Global vs Local Robustness

There are many situations where it is convenient and/or advisable and/or necessary, to limit the robustness analysis to a relatively small *neighborhood* of the uncertainty space under consideration. Hence the distinction between *global* and *local* robustness.

Informally, this distinction entails that a global robustness analysis is one that is conducted over the entire uncertainty space under consideration, whereas a local robustness analysis is one that is conducted over a relatively small *neighborhood* of the uncertainty space. Clearly, this distinction is reminiscent of the familiar distinction made in *optimization theory* between local and global *optimization.* It also reminds us of the distinctions between a local and a global *anesthetic,* a local and a global *economy,* local and global *weather*, local and global *news*, etc.

From a decision-making point of view, this distinction means that the choice of a robustness model, namely global versus local, should be compatible with, among other things, the type of uncertainty that we face, or postulate.

To explain this point, consider the situation depicted in Fig. 4.1, where the large rectangle represents the uncertainty space under consideration, denoted S, and the small white circle, denoted $\mathscr{B}$, represents a *neighborhood* of S.

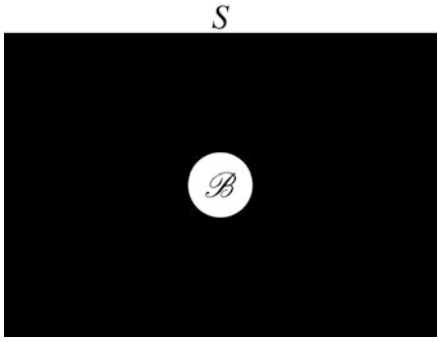


Fig. 4.1: An example of an uncertainty space S and one of its neighborhoods, $\mathscr{B}$

In this case, a *global* robustness analysis is one that is conducted over the *entire uncertainty space S*. In contrast, a robustness analysis that is conducted over a relatively *small neighborhood* of S, such as $\mathscr{B}$, is *local.*

From a robustness perspective, the point to note about the distinction between S and $\mathscr{B}$ is that $\mathscr{B}$ is not merely a small *subset* of S, it is a small *neighborhood* of S. The difference between "a small *subset* of" and "a small *neighborhood* of" is profound because by its very nature a *small neighborhood* of a large set gives only a *local* picture of the large set. On the other hand, in the context of Fig. 4.1, a small subset of S, say a dense *grid* over S, can give a very good indication of the variability of s over S.

Informally then, a global robustness analysis would be appropriate when we seek robustness against variations in the value of s over S, whereas a local robustness analysis would be appropriate when we seek robustness against variations in the value of s over a small *neighborhood* of S, say the neighborhood $\mathscr{B}$ depicted in Fig. 4.1.

And the implication is that as S denotes the set of all possible/plausible values of s, the use of a local robustness model needs to be *justified.* That is, we need to explain and justify on what grounds do we exclude from the robustness analysis all the possible/plausible values of s that are outside the neighborhood over which the local robustness analysis is conducted. Barring such a justification, the local analysis might be exposed to the valid criticism of lacking any logical foundation (see Sect. 4.7.3).

And with this as background, let us now examine how the distinction between a *local* and a *global* analysis is manifested in worst-case-based robustness models, notably maximin models.

In fact our objective is to make it clear that maximin models can, as a matter of principle, perform both local and global robustness analyses. To this end we need to recast, indeed extend, the definitions given to the sets $S(d), d \in S(d)$. Keeping in mind then that in Sect. 4.4 we let $S(d)$ denote the set of all the possible/plausible values of s that are generated by decision d, we now distinguish between three sets associated with decision d, namely:

$$S(d) := \text{ set of all possible/plausible states generated by decision } d. \tag{4.22}$$

$$\overline{S}(d) := \text{ subset of } S(d) \text{ used in the robustness analysis of decision } d. \tag{4.23}$$

$$\overline{\overline{S}}(d) := S(d) \setminus \overline{S}(d) \ \text{ (complement of } \overline{S}(d)). \tag{4.24}$$

Note that, by definition, the robustness analysis of decision d is confined to set $\overline{S}(d)$, meaning that it ignores all the states in $\overline{\overline{S}}(d)$.

Now consider the maximin model obtained from (4.3) by replacing $S(d)$ with $\overline{S}(d)$, namely consider this maximin model:

$$z^* := \max_{d \in D} \min_{s \in \overline{S}(d)} \{f(d,s) : constraints(d,s), \forall s \in \overline{S}(d)\}. \tag{4.25}$$

The implication is that in view of the above, if $\overline{S}(d) = S(d), \forall d \in D$, then this maximin model is a model of *global* robustness. And if for every $d \in D$ the set $\overline{S}(d)$ is a relatively small neighborhood of $S(d)$, then this maximin model is a model of *local* robustness. To illustrate, consider the following robustness model:

$$\rho(q|\tilde{u}) := \max_{\alpha \geq 0} \{\alpha : \text{stab-con}(q,u), \forall u \in U(\alpha, \tilde{u})\}\ ,\ q \in Q \tag{4.26}$$

where Q is some given set, $U(\alpha, \tilde{u})$ is a neighborhood of size α around $\tilde{u}$ associated with some uncertainty space $\mathscr{U}$, $\tilde{u} \in \mathscr{U}$ is a nominal value of u, and stab-con(q,u) is a list of constraints on (q,u) pairs. We refer to $\rho(q|\tilde{u})$ as the *radius of stability* of q at $\tilde{u}$.

In words: the *radius of stability* of q at $\tilde{u}$, denoted $\rho(q|\tilde{u})$, is equal to the size, α, of the largest neighborhood $U(\alpha,\tilde{u})$ around $\tilde{u}$ all whose elements satisfy the stability constraints listed in stab-con(q,u). This is illustrated in Fig. 4.2 where the large rectangle represents the uncertainty set $\mathscr{U}$, the gray area represents the values of u that violate the stability constraints, and the circles centered at $\tilde{u}$ represents neighborhoods around $\tilde{u}$.

Thus, the *radius of stability* of system q at $\tilde{u}$ is equal to the radius of the largest circle centered at $\tilde{u}$ that is contained in the white area.

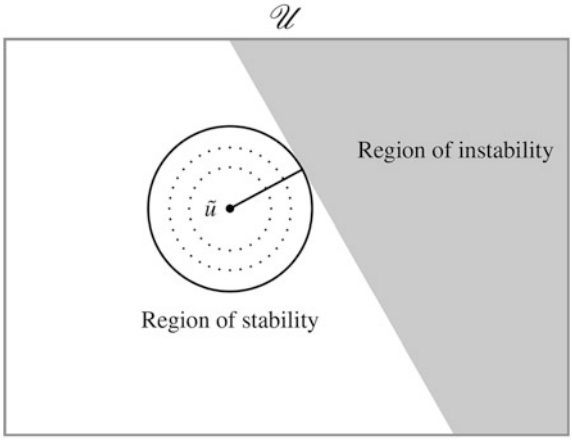


Fig. 4.2: Radius of stability of system q at $\tilde{u}$

Such models are used extensively in many fields to model the local stability and/or robustness of systems to perturbations in a nominal value of a parameter [23, 26, 50, 51].

In this context, seeking the most robust system entails seeking a system whose radius of stability is the largest, which amounts to solving this maximin problem:

$$\rho(\tilde{u}) := \max_{q\in Q}\ \max_{\alpha\geq 0}\ \{\alpha : \text{stab-con}(q,u), \forall u \in U(\alpha,\tilde{u})\} \tag{4.27}$$

$$= \max_{q\in Q,\alpha\geq 0}\ \{\alpha : \text{stab-con}(q,u), \forall u \in U(\alpha,\tilde{u})\}. \tag{4.28}$$

Observe then that *radius of stability* models are (local) maximin models. Specifically, the model specified by (4.26) is that instance of the maximin model that is specified by (4.25) yielded by setting $s = u$; $d = \alpha$; $D = [0,\infty)$; $S(d) = U(\alpha,\tilde{u})$; and $constraints(d,s) = \text{stab-con}(q,u)$.

The *global counterpart* of the radius of stability model specified in (4.26) is obtained by replacing the neighborhood $U(\alpha,\tilde{u})$ by a subset of the uncertainty space $\mathscr{U}$ and replacing α by a measure of the size of such a subset. The end result is this maximin model:

$$w(q|\tilde{u}) := \max_{V \subseteq \mathscr{U}} \{\#(V) : \text{stab-con}(q,u), \forall u \in V\}\ ,\ q \in Q \tag{4.29}$$

where $\#(V)$ denotes the "size" of set V. For example, if V consists of finitely many elements, we can let $\#(V)$ denote the *cardinality* of V.

In words, according to (4.29), the (global) robustness of q, denoted $w(q|\tilde{u})$, is equal to the size of the largest subset of $\mathscr{U}$ all whose elements satisfy the stability constraints.

In the idiom of the maximin paradigm, u denotes the state variable in this model, V the decision variable, and $\#(V)$ the payoff generated by V and u. Clearly, this is a state-free-payoff maximin model.

The situation depicted in Fig. 4.3 illustrates the difference between the radius of stability model and its global counterpart.

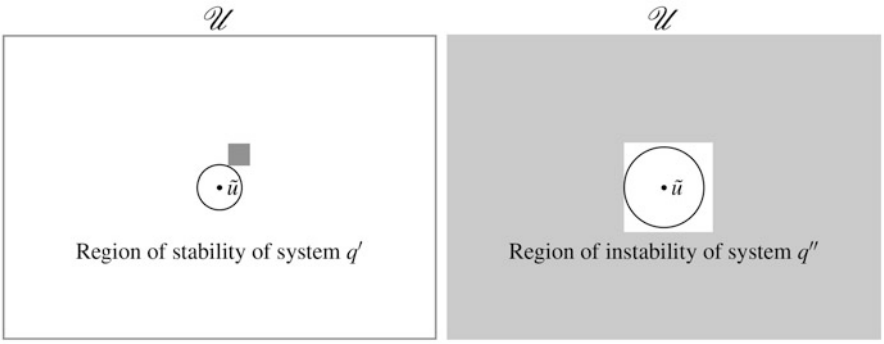


Fig. 4.3: Local and global robustness of two systems: q' and q''. The *white areas* represent regions of stability, the *gray areas* represent regions of instability and the *circles* represent the largest neighborhoods around $\tilde{u}$ that are contained in the respective regions of stability

The idea is then, as illustrated in Fig. 4.3, that if we measure the size of subsets of $\mathscr{U}$ by the "area" they take up, then the global robustness of a system would be equal to the size of the area covered by its *region of stability.* In this case, the robustness of the system would be equal to the size of the white area representing the region of stability of the system. Hence, according to this measure of robustness, in the context of Fig. 4.3, system q' is much more robust then system q''. In contrast, at $\tilde{u}$, the *radius of stability* of system q' is much smaller than the radius of stability of system q''. Hence, the implication is that, globally, decision d' is more robust than decision d'' whereas, locally at $\tilde{u}$, decision d'' is more robust than d'.

The point we want to make then is that, given its innate ability to provide a variety of measures of robustness, Wald's maximin paradigm proves an indispensable tool for robust decision-making.

4.7 A Robust Decision-Making Perspective

Before we take up the topic of *robust decision-making*, it is important to make it abundantly clear that in this discussion, the phrase *robust decision-making* refers to the wide-ranging, highly active multidisciplinary field of expertise that is concerned with a variety of theoretical, technical, and conceptual issues pertaining to robust decision-making, rather than to the specific methodology bearing this name that was developed at the *Rand Corporation* (see https://en.wikipedia.org/wiki/Robust_decision-making).

Thus, our main thrust in this section is to examine the role of the worst-case approach to uncertainty, more specifically that of Wald's maximin paradigm, as a tool of thought and as a practical instrument used in robust decision-making, notably for situations subject to a non-probabilistic uncertainty.

In fact, in certain disciplines, such as *robust optimization*, this paradigm dominates the scene. For this reason, it is instructive to examine how Wald's maximin paradigm came to play this role in this discipline.

4.7.1 Robust Optimization

One of the thorniest difficulties afflicting "conventional" optimization problems is that optimal solutions to such problems often prove sensitive, sometimes highly sensitive, to perturbations in the values of the problems' parameters. This fact renders the identification of "robust solutions" extremely important in *optimization theory*. To examine then how this issue arises in this field, let us consider the following abstract constrained optimization problem:

$$\text{Problem P}: \quad w^* := \max_{x \in X} \{g(x) : constraints(x)\} \tag{4.30}$$

where g is a real-valued function on some set X and $constraints(x)$ denotes a list of constraints on the decision variable x. We refer to g as the *objective function*. Let X^* denote the set of optimal solutions to this problem.

Now, suppose that both the objective function g and the constraints under consideration depend on a parameter, call it u, and let $\mathscr{U}$ denotes the set of all the values of u under consideration. We can thus deduce from Problem P the following *parametric optimization problem:*

$$\text{Problem P}(u): \quad w^*(u) := \max_{x \in X} \{g(x;u) : constraints(x;u)\}\ ,\ u \in \mathscr{U} \tag{4.31}$$

where the notation $(x;u)$ is used to highlight the fact that u is a *parameter* of the object under consideration. Let $X^*(u)$ denote the set of optimal solutions to this problem for the specified value of u. Note that for each $u \in \mathscr{U}$, Problem P(u) is a conventional constrained optimization problem.

Informally, a robust solution to this parametric optimization problem is a solution $x \in X$ that performs "well" (relatively to other solutions) with respect to the objective function and the constraints, against the variations in the value of parameter u. The following definition is therefore self-evident:

Definition. *An element of X, say x^*, is said to be a* SUPER-ROBUST *solution to the parametric problem (4.31) iff $x^* \in X^*(u), \forall u \in \mathscr{U}$, that is iff x^* is an optimal solution to Problem $P(u)$ for all $u \in \mathscr{U}$.*

There are of course situations where *super-robust* solutions exist, namely situations where the optimal solutions recovered for Problem P prove insensitive to the variation in the optimization problem's parameters. But such situations are the exception, rather than the rule.

The question therefore arising is this: what constitutes a "robust solution" to the parametric optimization problem specified by (4.31) in cases where no *super-robust* solutions exist?

If we adopt a *worst-case approach* to the variability of u over $\mathscr{U}$, then the answer to this question is as follows:

- A robust solution to the above parametric problem should satisfy the constraints under consideration for the worst u in $\mathscr{U}$ (whatever it is), hence for all $u \in \mathscr{U}$.
- A robust solution to this parametric problem should yield the best (largest over all $x \in X$ satisfying the worst-case constraints) worst (smallest over $u \in \mathscr{U}$) value of $g(x;u)$.

In short, a *worst-case approach* to the variability of u over $\mathscr{U}$ yields the following maximin counterpart for the above parametric problem:

$$\text{Problem RC:} \quad w^{(rc)} := \max_{x \in X} \min_{u \in \mathscr{U}} \{g(x;u) : constraints(x;u), \forall u \in \mathscr{U}\}. \tag{4.32}$$

In the literature on *robust optimization* (e.g. Ben-Tal et al. [5]) this maximin problem is called the *robust counterpart* problem. Any value of $x \in X$ that is an optimal solution to this maximin problem is regarded a *robust solution* to the parametric problem (4.31).

The idea of incorporating the maximin paradigm in the formulation of optimization problems so as to obtain solutions that are robust against variations in the values of the problem's parameters, goes back to at least the 1960s (e.g., Dorato and Drenick [15]). The phrase *robust optimization* became popular in the mid 1990s (e.g., Mulvey et al. [29]).

Interestingly, although in the broad literature on robust optimization the phrase *robust optimization* does not refer exclusively to maximin-based robustness (e.g., Mulvey et al. [29]), some scholars hold that robust optimization effectively boils down to the solution of maximin problems (e.g., Bertsimas et al. [8, pp. 465–466] and Ben-Tal et al. [7, p. 628]). This position, one need hardly point out, gives expression to the prominent role that the maximin paradigm has come to play in *robust optimization* in recent years.

In a nutshell, the element distinguishing the field of *robust optimization* from other fields where maximin models are used to obtain robust solutions to problems, is that in *robust optimization* the focus is on classical *mathematical programming problems* such as linear programming problems, quadratic programming problems, dynamic programming problems, and so on. For an overview of recent advances in robust optimization see [18].

4.7.2 Conservatism

For obvious reasons, decisions that are based on worst-case-analysis, compared to decisions that are based on other approaches to uncertainty, may be deemed *conservative*. Indeed, they might even be labeled *paranoid.* Because, given our individual and common experience, it seems safe to say that, while "worst-case scenarios" do occur, these would typically be "rare events". Clearly then, the worst-case approach to uncertainty is not based on evidence that the "worst-case scenario" is a good "estimate", or a good "approximation", of the "true" (unknown) scenario.

The question is then how should we understand the attribute "conservative" as it is applied to worst-case-analysis-based decisions considering the type of outcomes they yield.

The answer to that seems to be that this is very much a problem-oriented issue. The question as to whether a decision is "conservative" may depends on the problem one deals with, on the context in which it is dealt with, on the objectives one has, and so on.

The point to keep in mind here is that a decision that fares satisfactorily under a "worst-case scenario" would not necessarily fare satisfactorily under more "realistic" scenarios, particularly in the sense that it can turn out to be far too costly should a "realistic", rather than "worst-case scenario", be realized. Such a decision may well be deemed "conservative".

But, on the other hand, in cases where, for whatever reasons, the declared objective of an investigation, a project, and so on, is to identify decisions that perform well against "worst-case scenarios", then such decisions will obviously not be judged "conservative" at all.

And, as pointed out by Gabrel et al. [18, p. 472], the same applies for worst-case-based robustness:

> When uncertainty affects the *feasibility* of a solution, robust optimization seeks to obtain a solution that will be feasible for any realization taken by the unknown coefficients; however, complete protection from adverse realizations often comes at the expense of a severe deterioration in the objective. This extreme approach can be justified in some engineering applications of robustness, such as robust control theory, but is less advisable in operations research, where adverse events such as low customer demand do not produce the high-profile repercussions that engineering failures—such as a doomed satellite launch or a destroyed unmanned robot—can have.

As for the maximin paradigm, the charge of "conservatism" has been leveled at it ever since its introduction into *decision theory* in the 1950s. And to illustrate what is meant by this, consider the maximin problem associated with the following payoff table:

$f(d,s)$	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$	$s^{(4)}$	$s^{(5)}$	$s^{(6)}$	SL
$d^{(1)}$	2	2	2	2	2	1	1
$d^{(2)}$	999	999	999	999	999	0.99	0.99

(4.33)

In this case, the security level (SL) of decision $d^{(1)}$ is larger than the security level of decision $d^{(1)}$, hence the maximin paradigm, deems decision $d^{(1)}$ superior to (more robust than) decision $d^{(2)}$.

However, much as decision $d^{(1)}$ performs better than decision $d^{(2)}$ under the worst-case scenario (state), decision $d^{(2)}$ performs far better than decision $d^{(1)}$ under all other scenarios (states).

As we pointed out already, the verdict as to whether or not decisions yielded by the maximin paradigm are "conservative" is a problem-oriented issue. Thus, it may well apply in cases where one seeks decisions that perform well in relation to "realistic" rather than "worst-case" scenarios. But ...

Consider the point raised by Wasserman [49, Sect. 4]:

> The claim that minimax theory is driven by the worst case is a more substantial criticism. I happen to think worst case analysis is a good thing. I want an estimator that does reasonably well under any circumstances.

The inference therefore is that the deployment of the maximin paradigm as a framework for identifying decisions that perform well under "realistic" rather than "worst-case" scenarios, ought to be reasoned out carefully. It most certainly ought to be justified, regardless of whether the decisions yielded, for the case considered, are judged to be "conservative".

4.7.3 Irresponsible Decision-Making

On the face of it, rendering a worst-case analysis less "conservative" and more "realistic" seems to be straightforward. Indeed, all we need to do to this end is to exclude from the worst-case analysis "unrealistic" scenarios. Thus, in [18, p. 472], we read the following:

> To make the robust methodology appealing to business practitioners, robust optimization thus focuses on obtaining a solution that will be feasible for any realization taken by the unknown coefficients within a smaller, "realistic" set, called the *uncertainty set*, which is centered around the nominal values of the uncertain parameters. The goal becomes to optimize the objective, over the set of solutions that are feasible for all coefficient values in the uncertainty set. The specific choice of the set plays an important role in ensuring computational tractability of the robust problem and limiting deterioration of the objective at optimality, and must be thought through carefully by the decision maker.

To bring out the distinction between "realistic" and "unrealistic" scenarios (states) and some of the implications of this distinction, let

$$S = \text{ set of all possible/plausible scenarios (states).} \tag{4.34}$$

$$S^{(r)} = \text{ set of all the "realistic" scenarios in } S. \tag{4.35}$$

$$S^{(ur)} = \text{ set of all "unrealistic" scenarios in } S, \text{ namely } S^{(ur)} := S \setminus S^{(r)}. \tag{4.36}$$

Clearly, this distinction is similar to that made in Sect. 4.6.3 between the sets $S(d), \overline{S}(d)$ and $\overline{\overline{S}}(d)$ underlying the distinction between a *local* and a *global* analysis.

The implication is then that the proposition to deal with the "conservatism" issue by means of a worst-case analysis over $S^{(r)}$ rather than over S must be thoroughly justified. Namely, such a proposition must be accompanied by a cogent argument justifying the criteria used to distinguish between "realistic" and "unrealistic" scenarios.

Ben-Tal et al. [6] address a similar distinction between the *physically possible* values of the uncertainty parameter and the parameter's "normal range", arguing convincingly that a robustness model that effectively ignores all values of the parameter outside the "normal range" represents a '... somewhat "irresponsible" decision maker ...'.

Along the same line, in the article *Severe uncertainty and info-gap decision theory* [22, p. 609], point out that in the framework of *info-gap decision theory* there is a stark incongruity between the severity of the uncertainty under consideration and the local orientation of the robustness analysis, implying that, methodologically speaking, conclusions based on this theory may "... have no logical foundation...".

And in this vein, Sniedovich [39–42] argues that a failure to appreciate the distinction between local and global robustness and the ramifications of this distinction, especially with regard to situations subject to a severe uncertainty, may lead to "voodoo decision-making", namely to decision-making that is based on unrealistic, and/or misguided, and/or contradictory, and/or delusional assumptions.

From a "maximin perspective", this means that the worst-case analysis conducted by maximin models must be consistent with the quantification of the uncertainty under consideration and with the type of robustness sought by the decision maker. Thus, to prevent "irresponsible" decision-making, it is imperative to justify the exclusion of "unrealistic" scenarios from the robustness analysis and the criteria used to distinguish between "realistic" and "unrealistic" scenarios (states).

4.7.4 A Probabilistic Perspective on Worst-Case Analysis

One of the consequences of limiting a worst-case-based robustness analyses to "realistic" scenarios, is that such an analysis might produce results that would not be less risky than those produced by a *probabilistic* analysis that is conducted over a larger uncertainty space. Indeed, they might even be more risky. And to illustrate,

consider the following text comprising the abstract of the article *Risk analysis in robust control—making the case for probabilistic robust control* ([10]; emphasis added):

> This paper offers a critical view of the "worst-case" approach that is the cornerstone of robust control design. It is our contention that a blind acceptance of worst-case scenarios may lead to designs that are actually more dangerous than designs based on probabilistic techniques with a built-in risk factor. The real issue is one of modeling. If one accepts that no mathematical model of uncertainties is perfect then a probabilistic approach can lead to more reliable control even if it cannot guarantee stability for all possible cases. Our presentation is based on case analysts. We first establish that worst-case is not necessarily "all-encompassing". In fact, we show that for some uncertain control problems to have a conventional robust control solution it is necessary to make assumptions that leave out some feasible cases. Once we establish that point, we argue that **it is not uncommon for the risk of unaccounted cases in worst-case design to be greater than that of the accepted risk in a probabilistic approach.** With an example, we quantify the risks and show that worst-case can be significantly more risky. Finally, we argue that the deterministic worst-case analysis is not necessary more reliable than the probabilistic analysis.
>
> Chen et al. [10]

The idea here is that, while a probabilistic analysis would be conducted over the entire uncertainty space, a restricted worst-case analysis would leave out values of the uncertainty parameter that are outside a small subset of the uncertainty set.

Aside from that, it is also important to highlight another aspect of the relation between a deterministic worst-case analysis and a probabilistic analysis. This is the fact that certain probabilistic models have obvious deterministic worst-case equivalents, a fact that some scholars regard as *remarkable* (e.g., Elishakoff and Ohsaki [16, p. 12, p. 245]). The objective of the brief discussion below is to explain why this fact is anything but remarkable, observing that it is a direct implication of the relation between the fundamental probabilistic concepts *almost surely* and *surely*.

Recall then that certain probabilistic models are in fact "probabilistic in-name-only". That is, such models are not "genuine" probabilistic models but are rather *deterministic* models that are dressed (formulated) in a probabilistic garb. Differently put, they are "degenerate" probabilistic models that have obvious equivalent deterministic counterparts. Also, bear in mind that *surely* probabilistic events involving *constraints* have obvious worst-case deterministic counterparts.

For our purposes it suffices to examine a simple example. Consider then the following two constraints associated with random variable Y

Probabilistic constraint	Deterministic (worst-case) constraint
$P(h(Y) \in C) = 1$	$h(y) \in C, \forall y \in \mathbb{Y}$

(4.37)

where C is a subset of the real line, y denotes a realization of Y, $\mathbb{Y}$ denotes the set of possible realizations of Y, h is a real-valued function on $\mathbb{Y}$, and $P(E)$ denotes the probability of event E.

Next, recall the difference between *almost surely* and *surely* events: the former occur with probability 1 and the latter occur "for sure". To be precise, in the context of the above constraints,

$$\overbrace{h(y) \in C \ , \ \forall y \in \mathbb{Y}}^{\text{surely}} \longrightarrow \overbrace{P(h(Y) \in C) = 1}^{\text{almost surely}} \tag{4.38}$$

but not necessarily vice versa.

The point is, however, that there are many cases where *almost surely* events are also *surely* events, the implication being that in such cases we have

$$\overbrace{h(y) \in C \ , \ \forall y \in \mathbb{Y}}^{\text{surely}} \longleftrightarrow \overbrace{P(h(Y) \in C) = 1}^{\text{almost surely}} . \tag{4.39}$$

To illustrate, consider the case where $\mathbb{Y}$ and C are closed, bounded intervals of the real line, say $\mathbb{Y} = [\underline{y}, \overline{y}]$, and $C = [\underline{c}, \overline{c}]$; h is continuous on $\mathbb{Y}$, and the cumulative distribution function of Y is *strictly increasing* on $\mathbb{Y}$. Then clearly in this case $P(h(Y) \in C) = 1$ entails that $h(y) \in C \ , \ \forall y \in \mathbb{Y}$, hence (4.39) holds and C contains the set (interval) $h(\mathbb{Y}) := \{h(y) : y \in \mathbb{Y}\}$. Observe that this implies that the smallest interval C satisfying the probabilistic constraint $P(h(Y) \in C) = 1$ is $C^* = h(\mathbb{Y})$.

For similar reasons, subject to the above assumptions on h and Y, we have the following equivalence between a probabilistic constrained optimization problem and its deterministic worst-case counterpart:

Probabilistic Problem	Deterministic equivalent
$\underset{x \in X}{\text{opt}} \{f(x) : P(h(Y) \in H(x)) = 1\}$	$\underset{x \in X}{\text{opt}} \{f(x) : h(y) \in H(x), \forall y \in \mathbb{Y}\}$

(4.40)

where f is a real-valued function on some set X and $H(x)$ is a bounded interval of the real-line whose endpoints depend on x. These problems are equivalent in the sense that they possess the same optimal solutions. Observe that the deterministic equivalent problem is a state-free-payoff maximin or minimax problem, depending on whether opt = max or opt = min, respectively.

In sum, there is nothing remarkable, or surprising, about the fact that a deterministic worst-case analysis and a probabilistic analysis of constraints associated with *surely* events yield the same results.

4.8 Can Wald's Maximin Paradigm Save the World?

The heading of this closing section paraphrases the apparently tongue in cheek heading *Minimax Theory Saves the World* of a post on Prof. Larry Wasserman's website entitled *Normal Deviate* (see Wasserman [49, Sect. 4]), where we read the following:

> Minimax theory is the best thing in statistical machine learning—or the worst—depending on your point of view.

and

> **5. Minimaxity Gets the Last Word?**
> Minimax theory was important in statistics for a while then it sort of faded. But it had a revival, partly due to the work of Donoho et al.
>
> But what is really interesting is that I see more and more minimax analyses in ML outlets such as NIPS, JMLR etc. So perhaps the marketplace of ideas is telling us that minimax theory is gaining in importance?

where ML = machine learning, JMLR = Journal of Machine Learning Research, NIPS = Neural Information Processing Systems, Donoho et al. = Donoho et al. [14], Donoho and Johnstone [13].

An interesting perspective on Wald's maximin paradigm from the viewpoint of *machine learning* can be found in the article *Machine Wald* [32].

The provocative tone of the above heading and the claims associated with it raise a number of issues whose clarification should serve as a fitting ending to the discussion in this chapter.

To begin with, consider the observation regarding a revival in the engagement with this paradigm. Looking at the various literatures in the broad area of decision-making under uncertainty, it would appear that the growing interest in Wald's maximin paradigm in the past 20 years or so can in fact be attributed to a number of factors foremost of which is the development of *new algorithms* for the solution of large scale maximin problems [5, 8, 18].

But more than that, as much as the popularity of the paradigm may have waned and waxed over the years in various disciplines, the fact remains that for all the controversy surrounding it, Wald's maximin paradigm has retained its status as a major methodological and practical tool for decision-making under (a non-probabilistic) uncertainty in various areas of engineering, economics, management, operation research and so on.

It is important therefore to keep in mind that to correctly appreciate its scope of operation, its capabilities and its limitations, one must never lose sight of the fact that, methodologically speaking, the paradigm was not introduced as a general purpose tool for the treatment of non-probabilistic uncertain in a decision-making environment. Rather it was introduced as an *ad hoc* approach to such an uncertainty, where the only justification given [46–48] for its worst-case stance was that in the face of *complete ignorance* it is not unreasonable to ... assume the worst!

And to illustrate, consider the article *Statistical decision functions which minimize the maximum risk* where we read the following (Wald [47, p. 279]; emphasis added):

> A problem of statistical inference may be interpreted as a zero sum two person game as follows: Player 1 is Nature and Player 2 is the statistician. The variable r_1 is the parameter point θ the value of which is chosen by Nature. The variable r_2 is the statistical decision function $\omega(E)$ which is chosen by the statistician. The outcome $K[\theta, \omega(E)]$ of the game is the risk $r[\theta|\omega(E)]$ of the statistician. Clearly, the statistician wishes to **minimize** $r[\theta|\omega(E)]$. However, if the statistician is in **complete ignorance** as to Nature's choice, **it is perhaps not unreasonable** to base the theory of a proper choice of $\omega(E)$ on the assumption that Nature wants to **maximize** $r[\theta|\omega(E)]$.

The point is of course that even if we grant that the idea underlying the maximin paradigm is intuitive and comes naturally to us, grounding it on this thought because "*it is perhaps not unreasonable*", leaves it wide open to valid criticism, not to mention misuse and abuse. Because, strictly speaking, as pointed out by many scholars, there are no grounds for assuming that *Nature* is an adversarial opponent in this setup.

That said, it should be emphasized that the difficulty here is not that it is impossible to axiomatize Wald's maximin paradigm and to thereby ground it on a solid foundation, at least for a certain class of decision-making problems. Rather, the trouble is that such an axiomatization (e.g., Gilboa and Schmeider [19]) requires the presupposition of highly restrictive assumptions. Obviously, other decision theories encounter similar difficulties when it comes to axiomatization [17, 27]. But this is cold comfort.

Finally, we might add that, as intimated by Tintner [44, p. 24], apparently Wald himself was not fully satisfied with the paradigm and its conservative bent:

> Wald advocated the minimax principle in a tentative way and because of certain formal advantages. I am informed that he was still interested in finding a less conservative and more satisfactory principle for statistical inference.

And to sum it all up, methodologically and practically, Wald's maximin paradigm should be considered a tool that is suitable for decision-making in situations where it is reasonable (or advisable, or required) to adopt a worst-case approach to uncertainty, or variability. It is not, indeed, never was meant to be, a *panacea* for dealing with non-probabilistic uncertainty. Thus, as noted by Savage [37, pp. 578–579]:

> Studies of the minimax rule have been stimulating for statistics, and modifications and outgrowths of the rule may prove of great value, but those of us who, 12 or 13 years ago, hoped to find in this rule an almost universal answer to the dilemma posed by abstinence from Bayes' theorem have had to accept disappointment.

The inevitable conclusion therefore seems to be that, for all its long and distinguished service since 1939, Wald's maximin paradigm does not, indeed cannot "save the world". It continues, however, to offer an important tool of thought to decision makers dealing with uncertainty.

Hence, analysts and scholars who doubt the efficacy of this stalwart of robust decision-making should ask themselves whether *when in doubt* they would, or should, . . .

assume the worst!

References

1. Andronov, A.A., Pontriagin, L.S.: Robust systems. Dokl. Akad. Nauk SSSR **14**, 247–250 (1937) (in Russian)
2. Andronov, A.A., Vitt, A.A., Khaikin, S.E.: Theory of Oscillations. Nauka, Moscow (1959) (in Russian)

3. Barmish, B.R., Lagoa, C.M., Tempo, R.: Radially truncated uniform distributions for probabilistic robustness of control systems. In: Proceedings of the American Control Conference, Albuquerque, pp. 853–857 (1997)
4. Bell, K.J.: Some observations n the teaching and practices of pessimization. Heat Transfer Eng. **13**(1), 5–6 (1992)
5. Ben-Tal A., El Ghaoui, L., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
6. Ben-Tal, A., Golany, B., Shtern, S.: Robust multi-echelon multi-period inventory control. Eur. J. Oper. Res. **199**, 922–935 (2009)
7. Ben-Tal, A., Hazan, E., Koren, T., Mannor, S.: Oracle-based robust-optimization via online learning. Oper. Res. **63**(3), 628–638 (2015)
8. Bertsimas, D., Brown, D.B., Caramanis, C.: Theory and applications of robust optimization. SIAM Rev. **53**(3), 464–501 (2011)
9. Box, G.E., Andersen, S.L.: Permutation theory in the derivation of robust criteria and the study of departures from assumption. J. R. Stat. Soc. **17**(1), 1–34 (1955)
10. Chen, X., Aravena, J.L., Zhou, K.: Risk analysis in robust control—making the case for probabilistic robust control. In: Proceedings of the 2005 American Control Conference, Portland, vol. 3, pp. 1533–1538 (2005)
11. Chen, X., Zhou, K., Aravena, J.: Probabilistic robustness analysis—risks, complexity, and algorithms. SIAM J. Control. Optim. **47**(5), 2693–2723 (2008)
12. Corbett, J.O.: Uncertainty in risk analysis: an alternative approach through pessimisation. J. Radiol. Prot. **8**(2), 107–117 (1988)
13. Donoho, D.L., Johnstone, I.M.: Minimax estimation via wavelet shrinkage. Ann. Stat. **26**(3), 879–921 (1998)
14. Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D.: Wavelet shrinkage: asymptopia? J. R. Stat. Soc. B **57**(2), 301–369 (1995)
15. Dorato, P., Drenick, R.F.: Optimality, sensitivity, and game theory. In: Radanović, L. (ed.) Sensitivity Methods in Control Theory, pp. 78–102. Pergamon Press, New York (1966)
16. Elishakoff, I., Ohsaki, M.: Optimization and Anti-optimization of Structures Under Uncertainty. Imperial College Press, London (2010)
17. Ellsberg, D.: Risk, ambiguity and the Savage axioms. Q. J. Econ. **75**(4), 643–669 (1961)
18. Gabrel, V., Murat, C., Thiele, A.: Recent advances in robust optimization: an overview. Eur. J. Oper. Res. **235**(3), 471–483 (2014)
19. Gilboa, I., Schmeidler D.: Maximin expected utility with non-unique prior. J. Math. Econ. **18**, 141–153 (1989)
20. Guerrava-Vázquez, F., Rückmann, J.J.: Semi-infinite programming. In: Cochran, J.J. (ed.) Wiley Encyclopedia of Operations Research and Management Science. Wiley, New York (2010)
21. Hart, A., Roelofs, W., Crocker, J., Murray, A., Boatman, N., Hugo, S., Fitzpatrick, S., Flari, V.: Quantitative approaches to the risk assessment of GM crops. Defra Research Contract CPEC38. Central Science Laboratory, Sand Hutton (2007)
22. Hayes, K.R., Barry, S.C., Hosack, G.R., Peters, G.W.: Severe uncertainty and info-gap decision theory. Methods Ecol. Evol. **4**, 601–611 (2013)
23. Hindrichsen, D., Pritchard, A.J.: Stability radii of linear systems. Syst. Control Lett. **7**, 1–10 (1986)
24. IECRC: Chemical reaction engineering. Ind. Eng. Chem. Res. **58**(7), 15–17 (1966)
25. McKenzie, W.M.C.: Design of Structural Elements to Eurocodes. Palgrave, Basingstoke (2013)
26. Milne, W.E., Reynolds, R.R.: Fifth-order methods for the numerical solution of ordinary differential equations. J. ACM **9**(1), 64–70 (1962)
27. Milnor, J.: Games against nature. U.S. Air Force Project Rand, Research Memorandum RM-679 (1951)
28. Morgenstern, O.: Abraham Wald, 1902–1950. Econometrica **19**(4), 361–367 (1951)
29. Mulvey, J.M., Vanderbei, R.J., Zenios, S.A.: Robust Optimization of Large-Scale Systems. Oper. Res. **43**(2), 264–281 (1995)

30. Mutapcic, A., Boyd, S.: Cutting-set methods for robust convex optimization with pessimizing oracles. Optim. Methods Softw. **24**(3), 381–406 (2009)
31. Oskooi, A., Mutapcic, A., Noda, S., Joannopoulos, J.D., Boyd, S.P., Johnson, S.G.: Robust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveguides. Opt. Express **20**(19), 21558–21575 (2012)
32. Owhadi, H., Scovel, C.: Machine Wald (2015). arXiv:1508.02449, http://arxiv.org/abs/1508.02449
33. Pervozvanski, A.A.: Sensitivity, robustness and efficiency of adaptation. Int. J. Adapt. Control Signal Process. **6**, 183–191 (1992)
34. Rawls, J.: Theory of Justice. Belknap Press, Cambridge (1971)
35. Rustem, B., Howe, M.: Algorithms for Worst-case Design and Applications to Risk Management. Princeton University Press, Princeton (2002)
36. Savage, L.J.: The theory of statistical decision. J. Am. Stat. Assoc. **46**(253), 55–67 (1951)
37. Savage, L.J.: The foundation of statistics revisited. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley. **1**, 575–586 (1961)
38. Sheynin, O.B.: Origin of the theory of errors. Nature **211**, 1003–1004 (1966)
39. Sniedovich, M.: A bird's view of info-gap decision theory. J. Risk Finance **11**(3), 268–283 (2010)
40. Sniedovich, M.: Black swans, new Nostradamuses, voodoo decision theories and the science of decision-making under severe uncertainty. Int. Trans. Oper. Res. **19**(1–2), 253–281 (2012)
41. Sniedovich, M.: Fooled by local robustness. Risk Anal. **32**(10), 1630–1637 (2012)
42. Sniedovich, M.: The elephant in the rhetoric on info-gap decision theory. Ecol. Appl. **24**(1), 229–233 (2014)
43. Stein, O.: How to solve a semi-infinite optimization problem. Eur. J. Oper. Res. **223**, 312–320 (2012)
44. Tintner, G.: Abraham Wald's contributions to econometrics. Ann. Math. Stat. **23**(1), 21–28 (1952)
45. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944); Quotes from the third edition, 1953; sixth printing, 1955
46. Wald, A.: Contributions to the theory of statistical estimation and testing hypotheses. Ann. Math. Stat. **10**(4), 299–326 (1939)
47. Wald, A.: Statistical decision functions which minimize the maximum risk. Ann. Math. **46**(2), 265–280 (1945)
48. Wald, A.: Statistical Decision Functions. Wiley, New York (1950)
49. Wasserman, L.: Minimax Theory Saves The World (2012). https://normaldeviate.wordpress.com/2012/07/17/minimax-theory-saves-the-world/
50. Wilf, H.S.: Maximally stable numerical integration. J. Soc. Ind. Appl. Math. **8**(3), 537–540 (1960)
51. Zlobec, S.: Survey of input optimization. Optimization **18**, 309–348 (1987)

Chapter 5
The State of Robust Optimization

Seçil Sözüer and Aurélie C. Thiele

Abstract This survey presents a broad overview of the developments in robust optimization over the past 5 years, i.e., between 2011 and 2015. We highlight the advancement of knowledge both with respect to the theory of robust optimization and application areas. From a theoretical standpoint, we describe novel findings in static and multi-stage decision making, the connection with stochastic optimization, distributional robustness and robust nonlinear optimization. In terms of application areas, we consider inventory and logistics, finance, revenue management and health care. We conclude with guidelines for researchers interested in immunizing their problem against uncertainty.

5.1 Introduction

A classical assumption in mathematical programming is that the input data is perfectly known; however, in practice this is a rather rare situation and researchers have attempted to take data uncertainty into account since the seminal work of Charnes and Cooper [35] on chance-constrained programming. Unfortunately, many settings in today's fast-changing environments do not lend themselves to a probabilistic description of uncertainty. Robust optimization was first proposed in the early 1970s in order to provide a decision-making framework when probabilistic models are either unavailable or intractable, and has been the focus of significant research attention from the 1990s onwards.

S. Sözüer • A.C. Thiele (✉)
Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015, USA
e-mail: ses515@lehigh.edu; aut204@lehigh.edu

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_5

Robust optimization assumes that the uncertain data belongs to a convex and bounded set, called uncertainty set, and aims to find a solution that would remain feasible for all possible instances of the data parameters while achieving the best possible worst-case performance, as measured by the objective for the worst-case realization of the parameters. The specific choice of the set naturally plays an important role in terms of tractability and insightfulness of the optimal solution. Key to the tractability of robust optimization is the ability to optimize this worst-case criterion efficiently in presence of two conflicting imperatives: (1) a high level of robustness (protection against uncertainty) (2) the attainment of high-quality objective values (close to the objective of the nominal problem).

Soyster [107] took the first step toward the attainment of a robust optimization methodology in 1973. In order to find a solution immune to data uncertainty in linear programming, he injected the worst-case value of each uncertain parameter into the mathematical programming model; however, the model was deemed too conservative for practical implementation by many business practitioners. Then, in the late 1990s, significant progress in tackling the issue of over-conservatism was made by Ben-Tal and Nemirovski [14–16], El-Ghaoui and Lebret [45] and El-Ghaoui et al. [46]. These papers provided the foundation for modern robust optimization. (Note that the earlier paper of Mulvey et al. [91] uses a different concept also called robust optimization that builds upon the stochastic programming problem and optimizes a weighted combination of the traditional Stochastic Programming (SP) objective and a feasibility penalty function, which penalizes violations of the control constraints. This different definition for robust optimization will not be explored here.) The focus was mainly on constructing models more relevant to practitioners by controlling the degree of conservatism in uncertain linear problems with ellipsoidal uncertainty sets centered at the nominal value of the parameters. These problems were reformulated as second-order cone problems [15]. A drawback is that the resulting model is computationally less efficient than its nominal counterpart due to the added nonlinearity. This makes extensions to integer decision variables challenging from a computational standpoint.

In a milestone work, Bertsimas and Sim [25] investigated novel ways to decrease over-conservatism by tackling what they call the *Price of Robustness* using polyhedral uncertainty sets, which they connect to probabilities of constraint guarantees. Their approach offers full control on the level of conservatism for each constraint through a parameter called the budget of uncertainty that is adjusted by the decision maker. The interpretation of this budget of uncertainty is that it limits the number of uncertain parameters that can deviate from their nominal value. In this approach, the robust counterpart of a linear program remains linear, so that the robust model retains the advantages of a linear optimization model in terms of computational efficiency. Further, it can be readily generalized to discrete optimization, so that the robust counterpart of a integer linear program remains an integer linear problem.

For a comprehensive book treatment and survey on robust optimization, the reader is referred to Ben-Tal et al. [18] and Ben-Tal and Nemirovski [17], respectively. Also, Gorissen et al. [56] provide a practical guide to robust optimization that should be of significant interest to researchers attempting to immunize their prob-

lems against parameter ambiguity. Gabrel et al. [50] present an overview of recent advances in robust optimization between 2007 and 2012.

The present chapter focuses on studies indexed in Web of Science and published between 2011 and 2015 included, belonging to the area of Operations Research and Management Science, and having "robust" and "optimization" in their title. We narrowed the list of papers to over one hundred we deemed most significant by taking into account the research area, citation number, authors' track record in robust optimization and the journal's impact factor. This was necessarily a subjective process and some recent papers not listed here will certainly go on to have substantial impact on the field; however, we hope that this survey provides a good starting point into robust optimization today. Related book treatments and milestone works are also presented for reference. Papers are grouped by theme; within each theme they are listed in alphabetical order.

5.2 Theory of Robust Optimization

Since robust static (single-objective) linear programming is now well understood, current research efforts have mostly focused on (1) developing a stronger connection with stochastic optimization, (2) incorporating robust optimization to ambiguous probability distributions of random parameters rather than to ambiguous parameters of unknown but fixed value, (3) studying robust static nonlinear optimization, (4) considering multiple objective criteria, leading to the theory of robust Pareto efficiency, and (5) investigating robust dynamic decision-making. Note that Sniedovich [106] cautions against attempts to tackle severe uncertainty, characterized by a poor point estimate, a likelihood-free quantification of uncertainty and a large uncertainty space, using local robustness models based on the "radius of stability" concept.

5.2.1 Connection with Stochastic Optimization

In Stochastic Optimization, the uncertain data is assumed to be random. In the simplest case, these random parameters have a known probability distribution, while in more advanced settings, this distribution is only partially known. While robust optimization first emerged as a deterministic (worst-case) alternative to stochastic programming, each arising from different models of uncertainty, in recent years increasing numbers of researchers have strived to connect the robust optimization and stochastic optimization paradigms so that the models can be best tailored to the available information.

5.2.1.1 Foundational Work

The most important developments have led to a greater connection between the robust and stochastic optimization descriptions of uncertainty. They have been: (1) an argument that uncertainty sets, approached through robust optimization, should serve as the primitive for stochastic systems and (2) the design of safe tractable approximations of chance constraints to obtain guarantees of constraint violation and their Robust Counterpart representations and (3) a connection between linear problems with uncertain probabilities and uncertainty sets constructed as confidence sets using phi-divergences, with a size of the uncertainty set being controlled by the confidence level of the confidence set. Finally, a fourth work tackles robust nonlinear inequalities and thus develops tractable robust counterparts for new, previously unstudied classes of optimization problems.

Bandi and Bertsimas [8] investigate tractable stochastic analysis in high dimensions via robust optimization. They propose a new approach for stochastic systems based on robust optimization, to address the issue of computational tractability that arises when stochasticity is modeled using probabilities in areas such as queueing networks or multi-bidder auctions. Their framework relies on replacing the Kolmogorov axioms and the concept of random variables as primitives of probability theory, with uncertainty sets derived from some of the implications of probability theory like the central limit theorem. Performance analysis of stochastic systems in this new paradigm leads to linear, semidefinite or mixed integer optimization problems for which efficient algorithms capable of solving problems in high dimensions are available. Further, Bandi and Bertsimas [9] develop an optimal design framework for multi-item auctions based on robust optimization where they adopt an uncertainty set based model instead of using probability distributions.

Nemirovski [93] provides safe tractable approximations of chance constraints when data uncertainty is incorporated through randomly perturbed constraints. He reviews several simulation-based and simulation-free computationally tractable approximations of chance constrained convex programs, primarily, those found in chance constrained linear, conic quadratic and semidefinite programming, when the data is affinely parametrized by a random vector of partially known distribution. The models considered include Conditional Value-at-Risk and Bernstein approximations of the chance constraint. Robust counterpart representations of the approximations are also described.

Ben-Tal et al. [21] investigate robust linear optimization problems where the uncertain parameters with uncertainty regions defined by phi-divergences, which arise in settings involving moments of random variables and expected utility, and applications to asset pricing and the multi-item newsvendor problem. Phi-divergences refer to families of functions that measure "distance" between two vectors. The authors first derive confidence sets that are only asymptotically valid and then describe ways to improve the approximation by considering a modified statistic that uses correction parameters. They finally describe the robust counterpart with phi-divergence uncertainty and study its tractability. This is a special case of distributional robust

optimization, which we review in more details below. (The reader is also referred to Bayraksan and Love [11] for a tutorial on data-driven stochastic programming using phi-divergences.)

Finally, Ben-Tal et al. [22] present a model to formulate the robust counterpart of a nonlinear uncertain inequality that is concave in the uncertain parameters, using convex analysis and in particular Fenchel duality. Hence, robust models can be formulated for new classes of optimization models, for which tractable reformulations were not previously available. With respect to tractability, the authors further show that many robust counterparts can be written as linear, quadratic or conic quadratic constraints, or admit a self-concordant barrier function, which implies that the optimization problem can be solved in polynomial time.

5.2.1.2 Distributionally Robust Optimization and Chance Constraints

Ben-Tal et al. [19] consider chance constrained uncertain classification and investigate the problem of constructing robust classifiers when the training is plagued with uncertainty. They also discuss methodologies for classifying uncertain test data points and error measures for evaluating classifiers robust to uncertain data.

Dupacova and Kopa [42] consider stochastic programs whose set of feasible solutions depends on probability distributions that are not fully known, and adopt a contamination technique to study the robustness of results to perturbations on the probabilities. They suggest a robust efficiency test with respect to the second order stochastic dominance criterion.

With motivation drawn from data-driven decision making and sampling problems, Xu et al. [115] study the probabilistic interpretations of robust optimization by showing the connection between robust optimization and distributionally robust stochastic programming, and utilize this result to construct robust optimization formulations for sampled problems.

Zymler et al. [121] develop tractable approximations based on semidefinite programming for distributionally robust chance constraints where only the first- and second-order moments and support of the uncertain parameters are given. They investigate Worst-Case Conditional Value-at-Risk (CVaR) approximations and show that the approximation is tight for robust individual chance constraints with quadratic or concave constraint functions. For joint chance constraints, they show that the Worst-Case CVaR is provably tighter than two benchmark approximations. Further, a distributionally robust joint chance constrained optimization model for the case of the dynamic network design problem under demand uncertainty is developed by Sun et al. [108]. They propose an approach to approximate a joint chance-constrained cell transmission model based system optimal dynamic network design problem with only partial distributional information of uncertain demand.

Wiesemann et al. [113] consider Markov Decision Processes (MDP) with uncertain parameters when an observation history of the MDP is available. They derive a confidence region that contains the unknown parameters with a prespecified probability and obtain a policy that attains the best worst-case performance over this

confidence region, using the solution of conic programming problems of moderate size. Further, Wiesemann et al. [114] suggest a unifying framework for modeling and solving distributionally robust convex optimization problems based on standardized ambiguity sets that contain all distributions with prescribed conic representable confidence sets and encompass many ambiguity sets from the literature as special cases. They also model information about statistical indicators that have not yet been considered in the robust optimization literature, such as higher-order moments and the marginal median. The authors determine sharp conditions under which distributionally robust optimization problems based on their approach are computationally tractable, and tractable conservative approximations otherwise.

Alvarez-Miranda et al. [4] presents a note on the Bertsimas and Sim algorithm for robust combinatorial optimization problems with interval uncertainty, where they describe a method to solve fewer deterministic problems to obtain a robust solution. Long and Qi [85] investigate discrete optimization under the distributionally robust framework where they optimize the Entropic Value-at-Risk, a coherent risk measure that serves as a convex approximation of the chance constraint. They propose an approximation algorithm to solve the problem as a sequence of nominal problems and show in computational experiments that the number of nominal problems required is small for various distributional uncertainty sets.

Duzgun and Thiele [43] study 0-1 linear programming with uncertain objective coefficients using a safe tractable approximation of chance constraints, when the decision maker only knows the first two moments and the support of the random variables. They obtain a series of 0-1 linear programming problems parametrized by only one additional variable and show in numerical experiments that their model solves significantly faster than the benchmark robust model.

Zhen [119] investigates a variant of the task assignment problem under uncertainty based on stochastic programming and robust optimization. He develops both a stochastic programming model that tackles the issue of arbitrary probability distributions for the tasks' random workload requirements, and a robust optimization model which can cope with limited information about probability distributions.

Further, Duzgun and Thiele [44] bridge descriptions of uncertainty based on stochastic and robust optimization by considering multiple ranges for each uncertain parameter and setting the maximum number of parameters that can fall within each range, in a model reminiscent of histograms. The corresponding optimization problem can be reformulated in a tractable manner using the total unimodularity property of the uncertainty set and allows for a finer description of uncertainty while preserving tractability.

5.2.2 Nonlinear Optimization

We have already mentioned the work by Ben-Tal et al. [22], which presents a model to formulate the robust counterpart of a nonlinear uncertain inequality concave in the uncertain parameters. In this section, we list additional work pertaining to robust nonlinear optimization.

A specific case of nonlinear problems that are linear in the decision variables but convex in the uncertainty when the worst-case objective is to be maximized is investigated in Kawas and Thiele [76] in the context of portfolio management with uncertain continuously compounded rates of return. In that setting, exact and tractable robust counterparts can be derived. The authors extend their approach to short sales in [77], where they examine a class of non-convex robust optimization problems where the decision variables can be negative, leading to a non-convex problem in the uncertainty.

Ben-Tal and den Hertog [13] immunize conic quadratic optimization problems against ellipsoidal implementation errors. They prove that the robust counterpart of a convex quadratic constraint with ellipsoidal implementation error is equivalent to a system of conic quadratic constraints. They then extend the result to the case in which the uncertainty set is the intersection of two convex quadratic inequalities and show that the robust counterpart for this case is also equivalent to a system of conic quadratic constraints.

Doan et al. [41] build upon the fact that current successful methods for solving semidefinite programs are based on primal-dual interior-point methods and they approach robustness from an algorithmic perspective in order to address ill-conditioning and instability issues. Houska and Diehl [63] present a convex bilevel programming algorithm for the nonlinear min-max problems in semi-infinite programming. A conservative approximation strategy and optimality conditions are provided along with an analysis about strong global and locally quadratic convergence properties.

Poss [98] develops a robust combinatorial optimization model where the uncertain parameters belong to the image of multifunctions of the problem variables. A mixed-integer programming reformulation for the problem, based on the dualization technique is proposed since the feasibility set of the problem is non-convex. Jeyakumar and Li [69] focus on the trust-region problem, which minimizes a nonconvex quadratic function over a ball, and utilize the properties of the problems such as semi-definite linear programming relaxation (SDP-relaxation) and strong duality.

Finally, Suzuki et al. [109] investigate surrogate duality for robust nonlinear optimization and they prove surrogate duality theorems for robust quasiconvex optimization problems and surrogate min-max duality theorems for robust convex optimization problems. They provide necessary and sufficient constraint qualifications for surrogate duality and surrogate min-max duality, and give some examples at which such duality results are used effectively.

5.2.3 Multiple Objectives and Pareto Optimization

A large branch of Robust Optimization focuses on single-objective problems; however, multiple objectives are sometimes considered as well. Hu and Mehrotra [64] studies a family of models for multiexpert multicriteria decision making. Those

models utilize the concept of weight robustness in order to identify a (robust) Pareto decision that minimizes the worst-case weighted sum of objectives over a given weight region. The model is then extended to include ambiguity or randomness in the weight region as well as the objective functions. A multi-objective, multi-mode, multi-commodity, and multi-period stochastic robust optimization model is considered by Najafi et al. [92] where the purpose is to achieve the best possible emergency relief for earthquake response. Their method use hierarchical objective functions.

Fliege and Werner [49] consider general convex parametric multiobjective robust optimization problems under data uncertainty. They also present a characterization of the location of the robust Pareto frontier with respect to its nominal counterpart and illustrate their approach on a mean-variance problem. Robust optimization for interactive multiobjective programming with imprecise information is investigated by Hassanzadeh et al. [61] where there are clashing objectives and uncertainty occurs in both objective functions and constraints. They use an iterative procedure to capture the tradeoffs between the objectives.

Fang et al. [47] develop a multiobjective robust optimization model in order to enhance the performance and the robustness simultaneously. The multiobjective particle swarm optimization (MOPSO) algorithm is utilized for producing a set of non-dominated solutions over the entire Pareto space for a non-convex problem, which provides designers with more insightful information. Koebis [79] studied the relation between Scalar Robust Optimization and Unconstrained Multicriteria Optimization with a finite uncertainty set and showed that a unique solution of a robust optimization problem is Pareto optimal for the unconstrained optimization problem.

Iancu and Trichakis [66] incorporate Pareto efficiency to robust linear optimization problems and they present a characterization of Pareto robustly optimal solutions. Specifically, they argue that the classical RO paradigm may not produce solutions that possess the associated property of Pareto optimality, leading to potential inefficiencies and they propose practical methods that generate Pareto robustly optimal solutions by solving optimization problems that are of the same complexity as the underlying robust problems. Their numerical experiments are drawn not only from portfolio optimization—the best-known application area for Pareto optimal solutions—but also inventory management and project management. Hu and Mehrotra [65] consider robust decision making over a set of random targets or risk-averse utilities. In their setting, the random target has a concave cumulative distribution function or a risk-averse manager's utility is concave. Finally, Tong and Wu [111] investigate robust reward-risk ratio optimization models based on the positive homogenous and concave/convex measures of reward and risk.

5.2.4 Multi-Stage Decision-Making

While the main focus of robust optimization was static decision making when it was first investigated in the 1990s (following Soyster's 1973 work), multi-stage robust decision making has garnered substantial attention in recent years. In this setting,

uncertainty is revealed in stages and the decision maker adjusts his strategy based on the new information. The ability to take recourse action also allows the decision maker to tackle over-conservatism issues that affect static robust optimization when applied over multiple time periods. Delage and Iancu [40] provide an excellent tutorial on robust multi-stage decision-making.

5.2.4.1 Two Stages

Due to the difficulty inherent in multiple stages, many works have focused on two-stage robust optimization. The most notable works in this category are Bertsimas et al. [28], Hanasusanto et al. [60] and Ben-Tal et al. [23].

Bertsimas et al. [28] analyze the performance of static solutions for two-stage adjustable robust linear optimization problems with uncertain constraint and objective coefficients. They show that for a fairly general class of uncertainty sets, a static solution is optimal for two-stage adjustable robust linear optimization, which is quite counter-intuitive since static policies are generally believed to be conservative. Further, they develop a tight characterization of the adaptivity gap when no static solution is optimal. Their results lead to new geometric intuition about the performance of static robust solutions for adjustable robust problems, based on a certain transformation of the uncertainty set which helps highlight properties of the set when static robust policies do not perform well. Hence, the paper provides guidance in selecting the uncertainty set such that the adjustable robust problem can be well approximated by a static solution.

Hanasusanto et al. [60] extends the robust optimization methodology to problems with integer recourse, by approximating two-stage robust binary programs by their corresponding K-adaptability problems, in which the decision maker pre-commits in the first stage to K second-stage policies and implements the best of these policies once the uncertain parameters are realized. The authors study the quality of their approximation and the computational complexity of the K-adaptability problem. Further, they propose two mixed-integer linear programming reformulations that can be solved with off-the-shelf software.

Ben-Tal et al. [23] develop a method for approximately solving a robust optimization problem using tools from online convex optimization, where at every stage a standard (nonrobust) optimization program is solved. They find an approximate robust solution using a number of calls to an oracle that solves the original (nonrobust) problem that is inversely proportional to the square of the target accuracy. Their approach yields significant computational benefits when finding the exact solution of the robust problem is a NP-hard problem, for instance in the case of robust support vector machine with an ellipsoidal uncertainty set.

Additional work includes the following. Minoux [89] introduces a new subclass of polynomially solvable two-stage robust optimization problems with uncertainty on the right-hand side coefficients. Remli and Rekik [101] investigate the problem of combinatorial auctions in transportation services under uncertain shipment volumes and develop a two-stage robust formulation where they use a constraint

generation algorithm. Chan et al. [33] propose a computationally tractable and dynamic multi-stage decision methodology that can hedge against uncertainty by utilizing information from the previous stage iteratively, with an application to IMRT (intensity-modulated radiation therapy) treatment planning for lung cancer. Bo and Zhao [118] solve two-stage robust optimization problems by developing a column-and-constraint generation algorithm and compare their approach with the existing Benders-style cutting plane methods.

5.2.4.2 Optimal and Approximate Policies

We have already mentioned Bertsimas et al. [28], where the authors investigate the performance of static policies in two-stage robust linear optimization. Further, Bertsimas et al. [27] analyze the effect of geometric properties of uncertainty sets, such as symmetry, in the power of finite adaptability in multi-stage stochastic and adaptive optimization. They investigate finitely adaptable solutions, which generalize the notion of static robust solutions in the sense that a small set of solutions is specified for each stage and the solution policy implements the best solution from the set, depending on the realization of the uncertain parameters in past stages. In particular, they show that a class of finitely adaptable solutions is a good approximation for both the multistage stochastic and the adaptive optimization problem.

Kuhn et al. [80] consider primal and dual linear decision rule policies in stochastic and robust programming, and compute the loss of optimality due to this policy. They show that both approximate problems are equivalent to tractable linear or semidefinite programs of moderate sizes. Shapiro [104] considers the adjustable robust approach to multistage optimization, derives related dynamic programming equations and connects the problem to risk-averse stochastic programming. He also shows that, as in the risk-neutral case, a basestock policy is optimal.

Supermodularity and affine policies in a particular class of dynamic robust optimization problems are investigated by Iancu et al. [67]. They aim to provide a connection between dynamic programming and decision rules, and solve tractable convex optimization problems. Bertsimas and Goyal [24] consider adjustable robust versions of convex optimization problems where the constraints and objectives are uncertain and they show that a static robust solution yields a good approximation for these problems under general assumptions.

5.3 Application Areas of Robust Optimization

5.3.1 Classical Logistics Problems

5.3.1.1 Newsvendor Problem

The newsvendor problem is the building block of modern inventory theory. While robust newsvendor problems were first studied long before the time window for

publication of interest in this review, they continue to be the focus of significant research. Jiang et al. [71] consider robust newsvendor competition under asymmetric information about future demand realizations. They devise an approach based on absolute regret minimization and derive closed-form expressions for the robust optimization Nash equilibrium solution for a game with an arbitrary number of players. Qiu et al. [100] investigate the robust inventory decision-making problem faced by risk-averse managers with incomplete demand information with ellipsoidal uncertainty in a newsvendor setting. Three basic models are developed: expected profit maximization, Conditional Value-at-Risk (or CVaR)-based profit maximization, and a combination of these two.

Finally, Hanasusanto et al. [59] consider multi-item newsvendor problems from a distributional robust optimization perspective when the demand distributions are multimodal. The products considered are subject to fashion trends that are not fully grasped at the time when orders are placed. Spatially separated clusters of probability mass lack a complete description. The decision-maker minimizes the worst-case risk of the order portfolio over all distributions compatible with the modality information. The authors show the robust problem admits a conservative, tractable approximation using quadratic decision rules, which achieves a high level of accuracy in numerical tests.

5.3.1.2 Combinatorial Optimization Problems

Remli and Rekik [101] study the robust winner determination problem for combinatorial auctions in transportation services when shipment volumes are uncertain and propose a two-stage robust formulation solved using a constraint generation algorithm.

Poss [98] extends the Bertsimas-and-Sim model for robust combinatorial optimization using variable budgeted uncertainty, which is less conservative than (traditional) budget of uncertainty for vectors with few non-zero components. The author uses a mixed-integer programming reformulation for the problem and compare his approach with that of Bertsimas and Sim on the robust knapsack problem, where variable budgeted uncertainty achieves a reduction of the price of robustness by an average of 18 %.

Chassein and Goerigk [36] propose a new bound for the midpoint solution in minmax regret optimization, which evaluates a solution against the respective optimum objective value in each scenario and aims to find robust solutions that achieves the lowest worst-case difference between the two. Heuristics with performance guarantees have potentially great value in this context because most polynomially solvable optimization problems have strongly NP-hard minmax regret counterparts. One of these approximations is the midpoint solution, obtained when the decision maker approximates the uncertain parameters by the average of their lower and upper bound and solves that problem. They derive an instance-dependent performance guarantee for the midpoint solution of at most 2 and apply their methodology to the robust shortest path problem.

5.3.1.3 Scheduling

Robust berth scheduling at marine container terminals where vessel arrival and handling times are uncertain is studied by Golias et al. in [55]. They propose a bi-objective optimization problem and a heuristic algorithm, and test the results using simulation.

Varas et al. [112] focus on production scheduling for a sawmill where the uncertainty arises from the supply of logs and the finished product orders. Using a two-stage adaptive robust optimization approach, Lima et al. [84] investigate weekly self-scheduling, forward contracting, and pool involvement for an electricity producer operating a mixed power generation station.

Che et al. [37] study the cyclic hoist scheduling problem with processing time window constraints. The uncertainty comes from the perturbations or variations of certain degree in the hoist transportation times. The authors propose a method to measure the robustness of a cyclic hoist schedule and develop a bi-objective mixed integer linear programming model to optimize cycle time and robustness.

5.3.2 Facility Location

Facility location is concerned with the optimal placement of facilities to minimize the design and transportation costs while considering factors such as customer satisfaction, covering/serving a certain area, or avoiding placing hazardous materials near housing. Baron et al. [10] applied robust optimization to a capacitated multi-period fixed-charge network location problem in a network under uncertain demand over multiple periods. Their goal is to determine the number of facilities, their location and capacities, as well as the production amount and allocation of demand to facilities.

Another network design problem has been studied by Li et al. [83], for the planning of network infrastructure such as roads, pipelines and telecommunication systems. Uncertainty originates from the demand, and maintenance related issues such as operating costs, degradation rates. They propose an efficient and tractable approach for finding robust optimum solutions to linear and quadratic programming problems with interval uncertainty using a worst case analysis.

Robust hub location problems are studied in Alumur et al. [2] where the uncertainty arises due to the set-up costs for the hubs and the demands to be transported between the nodes. The authors analyze the changes in the solutions driven by the different sources of uncertainty when considered either in isolation or in combination.

Guelpinar et al. [58] consider a stochastic facility location problem in which multiple capacitated facilities serve customers with a single product, given uncertain customer demand and a constraint on the stock-out probability. Robust optimization strategies for facility location appear to have better worst-case performance than non-robust strategies.

Gabrel et al. [51] investigate a robust version of the location transportation problem with an uncertain demand using a two-stage formulation. The resulting robust formulation is a convex (nonlinear) program, and the authors apply a cutting plane algorithm in order to solve the problem exactly. Finally, Ghezavati et al. [53] investigate the optimization of reliability for a hierarchical facility location problem under disaster relief situations by a chance-constrained programming, with the aim of rapidly bringing the appropriate emergency supplies to the affected villages.

5.3.3 Supply Chain Management

Supply chain problems deal with the management of the flow of goods and services from the producer to the customer. It includes the movement and storage of raw materials, work-in-process inventory, and finished goods from point of origin to point of consumption in a way that ensures good service level and high profit. There exists uncertainty in many parts of a supply chain especially due to demand uncertainty.

A production planning problem in small-size furniture companies has been studied by Alem et al. [1]. They utilized robust optimization tools to derive robust combined lot-sizing and cutting-stock models when production costs and product demands are uncertainty. Their motivation to adopt robust optimization instead of two-stage stochastic programming was the absence of an explicit probabilistic description of the input data and the incentive of not having to deal with a large number of scenarios in robust optimization.

Aouam and Brahimi [6] considered an integrated production planning problem and order acceptance decisions under demand uncertainty. Orders/customers are classified into classes with respect to the marginal revenue, quantity they are willing to buy and reliability assessment. Their model provides flexibility to decide on the fraction of demand to be satisfied from each customer class and consider production-related constraints as well as factors such as congestion on production lead times. An order acceptance strategy allows the decision maker to maintain an appropriate level of utilization.

Schoenlein et al. [103] investigate the measurement and optimization of the robust stability of multiclass queueing networks with an application to dynamic supply chains. Stability of these networks implies that the total number of customers in the network remains bounded over time. The authors rely on fluid network analysis to quantify robustness using a single number, called the stability radius.

Qiu and Shang [99] study robust multi-period inventory decisions for risk-averse managers with partial demand distribution information for products with a short life cycle. The three inventory models we developed aim respectively to maximize expected profit, maximize conditional value-at-risk-based profit, and balance between the two objectives where the corresponding robust counterparts are presented.

Ashayeri et al. [7] consider a supply chain where a company faces bankruptcy to fulfill its debt obligation with limited financial resources. The uncertainty arises from demands and exchange rates. They formulate a MIP model with specific downsizing features, which maximizes the utilization of resources through a combined operation of demand selection and production assets reallocation. A pulp production planning and supply chain management has been studied in Carlsson et al. [32]. They utilize a robust optimization approach to handle the demand uncertainty and to establish a distribution plan, together with related inventory management. In this setup, they observe that there is no need for explicit safety stock levels and they achieve higher profit. Kawas et al. [78] study a game-theoretic setup of a production planning problem under uncertainty in which a company is exposed to the risk of failing authoritative inspections due to non-compliance with enforced regulations.

Finally, Kang et al. [74] investigate distribution-dependent robust linear optimization with applications to inventory control where every element of the constraint matrix is subject to uncertainty and is modeled as a random variable with a bounded support.

5.3.4 Industry-Specific Applications

In this section, we reference papers on three industry-specific logistics-driven applications that have received substantial attention in the robust optimization literature.

In *warehouse management*, Ang et al. [5] propose a robust storage assignment approach in unit-load warehouses facing variable supply and uncertain demand in a multi-period setting. They assume a factor-based demand model and minimize the worst-case expected total travel in the warehouse with distributional ambiguity of demand.

In *train timetabling operations*, Cacchiani et al. [30] focus on Lagrangian heuristics the application of train time-tabling. Galli [52] describes the models and algorithms that arise from implementing recoverable robust optimization to train platforming and rolling stock planning, where the concept of recoverable robustness has been defined in Liebchen et al. A survey of nominal and robust train timetabling problems in its nominal and robust versions is presented in Cacchiani and Toth [29].

In the sawmill planning problem, in addition to previously-mentioned Varas et al. [112], which focuses on production scheduling for a sawmill where the uncertainty arises from the supply of logs and the finished product orders, Alvarez and Vera [3] consider a related formulation where variability affects the yield coefficients related to the cutting patterns used. Finally, Ide et al. [68] investigate an application of deterministic and robust optimization in the wood cutting industry with the goal of attaining resource efficiency.

5.3.5 Finance

5.3.5.1 General Portfolio Problems

Robust portfolio optimization is studied by Ye et al. [117] in the context of a Markowitz mean-variance model with uncertainty on mean and covariance matrix. They formulate the robust problem as a second-order cone programming problem and show in computational experiments that the portfolios generated by the robust model are not as sensitive to input errors as the ones given by the classical model.

Nguyen and Lo [94] develop robust portfolio optimization models based on investors' rankings of the assets instead of estimates of their parameters such as expected returns, when the ranking is subject to uncertainty. They solve a robust ranking problem using a constraint generation scheme. Marzban et al. [87] study a multiperiod robust optimization model including stocks and American style options. The decision maker selects the level of robustness through the length and the type of the uncertainty set.

5.3.5.2 Risk Measures

Chen et al. [38] considers robust portfolio problems where expected utility is maximized under ambiguous distributions of the investment return, while Moon and Yao [90] investigate robust portfolio management when absolute deviation from the mean is used as a risk measure, leading to a linear programming problem. The authors test the robust strategies on real market data and discuss performance of the robust optimization model based on financial elasticity, standard deviation, and market condition such as growth, steady state, and decline in trend.

Fertis et al. [48] propose the concept of robust risk measure, defined as the worst possible of predefined risks when each among a set of given probability measures is likely to occur. In particular, they introduce a robust version of CVaR and of entropy-based risk measures, and show how to compute and optimize the Robust CVaR using convex duality methods.

Kakouris and Rustem [73] consider robust portfolio optimization with copulas, where copulas are used to describe the dependence between random variables. They provide the copula formulation of the CVaR of a portfolio and extend their approach to Worst Case CVaR (WCVaR) though the use of rival copulas exploiting a variety of dependence structures.

Kapsos et al. [75] investigate the worst-case robust Omega ratio, where the Omega ratio is a performance measure addressing the shortcomings of the Sharpe ratio and is defined as the probability weighted ratio of gains versus losses for some threshold return target. The authors investigate the problem arising from the probability distribution of the asset returns being only partially known and show that the problem remains tractable for three types of uncertainty.

In the most recent body of work, Lagos et al. [81] analyzes the characterizations of the robust uncertainty sets related to coherent and distortion risk measures and

aim to mitigate estimation errors of the Conditional Value-at-Risk. Maillet et al. [86] investigate global minimum variance portfolio optimization under some model risk based on a robust regression-based approach. The robust portfolio corresponds to the global minimum variance portfolio in the worst-case scenario and it provides protection against errors in the reference sample covariance matrix. Finally, Bertsimas and Takeda [26] study optimization over coherent risk measures and non-convexities where the relation between coherent risk measures and uncertainty sets of robust optimization is taken into consideration.

5.3.6 Machine Learning and Statistics

The incorporation of Machine Learning and Robust Optimization is a growing field. The reader is referred to Caramanis et al. [31] for an overview of robust optimization in machine learning. Ben-Tal et al. [19] focus on the problem of constructing robust classifiers when the training is subject to uncertainty. The problem is formulated as a chance-constrained program that is relaxed utilizing Bernstein's approximation to yield a second-order cone problem whose solution is guaranteed to be feasible for the original problem. Xu et al. [116] study robust principal component analysis in the presence of contaminated data.

Ozmen et al. [96] utilize Conic Multivariate Adaptive Regression Splines (CMARS) for generalizing the model identification problem including the existence of uncertainty with the aim to increase the trustworthiness of the solution in case of data perturbation. Beliakov and Kelarev [12] study global non-smooth optimization in robust multivariate regression where the objective is non-smooth, non-convex and expensive to calculate. They analyze the numerical performance of several derivative-free optimization algorithms with the aim of computing robust multivariate estimators.

Support vector machine (SVM) classifiers with uncertain knowledge sets via robust optimization are studied by Jeyakumar et al. [70]. They show how data uncertainty in knowledge sets can be handled in SVM classification and provide knowledge-based SVM classifiers with uncertain knowledge sets using convex quadratic optimization duality.

5.3.7 Energy Systems

Another area that has seen significant growth recently is robust optimization in energy. An application of robust optimization to renewable energy, specifically wind energy, is investigated in Jiang et al. [72], with the objective of providing a robust unit commitment schedule for the thermal generators in the day-ahead market that minimizes the total cost under wind output uncertainty.

Classen et al. [39] study a robust optimization model and cutting planes for the planning of energy-efficient wireless networks under demand uncertainty where they apply three different cutting plane methods. Goryashko and Nemirovski [57] study robust energy cost optimization of a water distribution system with uncertain demand with the aim to optimize daily operation of pumping stations based on the concept of Affinely Adjustable Robust Optimization.

Lima [84] works on weekly self-scheduling, forward contracting, and pool involvement for an electricity producer under three different scenarios, corresponding to electricity price forecasts. Sauma et al. [102] adopt a robust optimization approach to assess the effect of delays in the connection-to-the-grid time of new generation power plants over transmission expansion planning where the uncertainty arises from construction times of new power plants. Finally, Zugno and Conejo [120] work on the energy and reserve dispatch in electricity markets where they cast the problem as an adaptive robust optimization problem instead of a stochastic programming problem due to computational efficiency issues.

5.3.8 Public Good

The public good applications aim to improve the health, safety and well-being of the general public. Two main fields are humanitarian relief and health care applications. Examples include determining treatment plans in a hospital, patient transportation among hospitals, patient-doctor scheduling and constructing emergency evacuation routes during a disaster (fire, tsunami, earthquake).

5.3.8.1 Humanitarian Logistics/Emergency Logistics Planning

After a disaster occurs, humanitarian and state organizations gather resources and staff to serve a community's needs in an efficient way. Robust optimization has great relevance in humanitarian relief supply chains since we face data uncertainty during disasters.

Ben-Tal et al. [20] investigate a robust logistics plan generation methodology that can hedge against demand uncertainty. They study the dynamic emergency response assignment and evacuation traffic flow problems. They apply an affinely adjustable robust counterpart approach in order to provide better emergency logistics plans. A multi-objective robust optimization model for logistics planning during earthquake is proposed in Najafi et al. [92]. This paper propose a multi-objective, multi-mode, multi-commodity, and multi-period stochastic model to manage the scarce sources efficiently and they ensure that the distribution plan performs well under the various situations due to robustness.

Tajik et al. [110] adopt a robust optimization approach for the pollution routing problem with pickup and delivery under uncertain data where the aim is to reduce fuel consumption and decrease green house gases emission due to their harmful

effects on environment and human health. Their study addresses a new time window pickup-delivery pollution routing problem (TWPDPRP) to deal with uncertain input data.

The most recent developments in robust humanitarian logistics are the following. Lassiter et al. [82] consider the flexible allocation of the workforce after a disaster in order take into account changing (uncertain) needs and volunteer preferences. They use robust optimization to handle the uncertainty in task demands and derive Pareto optimality and allocation decisions for any level of conservativeness. Ghezavati et al. [53] investigate a hierarchical facility location problem under disaster relief situations where robust optimization and chance-constrained programming are applied. Shishebori and Babadi [105] design a robust and reliable medical services network under uncertain environment and system disruptions. Finally, Paul and Wang [97] study the United States Department of Agriculture food aid bid allocations, which aims at providing food aid annually in response to global emergencies and famine.

5.3.8.2 Health Care Applications

Chan et al. [33] consider an adaptive robust optimization approach to IMRT (intensity-modulated radiation therapy) treatment planning for lung cancer. They propose a computationally tractable and dynamic multi-stage decision methodology that can hedge against uncertainty by utilizing the information from the previous stage iteratively. Nha et al. [95] develops a new robust design optimization procedure based on a lexicographical dynamic goal programming approach for implementing time-series based multi-responses for drug formulations in the pharmaceutical industry.

Holte and Mannino [62] study the problem of allocating scarce resources such as operating rooms or medical staff to medical staff when the exact number of patients for each specialty is uncertain and when the allocation is defined over a short period of time such as a week and subsequently repeated over the time horizon. They adopt an adjustable optimization approach and develop a row and column generation algorithm to solve it efficiently.

Chan et al. [34] consider a robust-CVaR optimization approach with application to breast cancer therapy where the loss distribution is dependent on the state of an underlying system and the fraction of time spent in each state is uncertain. Finally, Meng et al. [88] investigate a robust optimization model for managing elective admission in a public hospital, given the priority of emergency patients over elective ones. They propose an optimized budget of variation approach that maximizes the level of uncertainty the admission system can withstand without violating the expected bed shortfall constraint and solve the robust optimization model by deriving a second order conic programming counterpart of the model.

5.4 Conclusions and Guidelines for Implementation

We have provided an overview of recent developments in robust optimization over the past 5 years. As robust optimization is now about 20 years old, it has become a well-established tool to address decision-making under uncertainty but also remains a thriving research area. We remind the reader of the practical guide to implementing robust optimization provided in Gorissen et al. [56]. The researcher interested in implementing robust optimization faces several modeling choices, which will impact the structure of the robust problem, its tractability and the insights the decision maker can gain into the optimal solution.

First, should the uncertainty be on the problem parameters themselves (leading to the classical robust optimization paradigm) or their underlying probabilistic distributions (yielding distributionally robust optimization or DRO)? DRO is particularly suitable if the stochastic programming version of the problem is tractable and the decision maker feels confident that he knows specific attributes of the family of probability distributions, such as their first two moments. If the SP version of the problem suffers from tractability issues, then adding robustness to that formulation will make the problem at least as computationally demanding; hence, it will then be more promising to apply robust optimization to the ambiguous parameters.

Second, what is the type of uncertainty set most suitable for the problem at hand? When the uncertainty is on the ambiguous parameters, the decision maker can then either use polyhedral uncertainty sets, which do not change the complexity of the mathematical programming problems considered but lead to additional constraints and variables in the tractable reformulation, or ellipsoidal uncertainty sets, which do not require any new variable or constraint but introduce non-linearities. When some decision variables are integer, polyhedral uncertainty sets thus seem particularly suitable. When the uncertainty is on the probability distributions, the uncertainty set may for instance incorporate knowledge of support, mean, covariance, directional deviations in the manner of Goh and Sim [54].

Third, is it possible to take corrective action after part of the uncertainty is revealed? If yes, adaptive or adjustable robust optimization will be advisable to address potential over-conservatism issues and lead to decision rules that are easy to implement in practice. The choice of those decision rules and the fine-tuning of their parameters have implications on computational tractability, closeness to optimality and insightfulness of the optimal solution.

In today's fast-changing environment, robust optimization presents an appealing framework that is both intuitive and lends itself to computationally tractable reformulations that either are exact or approximations documented in numerical experiments to perform well against benchmarks. RO is hence expected to keep increasing in relevance and importance in the arsenal of decision making tools of the operations research professional. In the future, researchers are likely to continue investigating improved approaches to multi-stage optimization, and to further connect RO with SP in order to provide an integrated approach to decision-making under uncertainty. Cutting-edge areas of interest include, but are not limited to, complex problems

such as adversarial risk analysis, policy design, performance evaluation, optimization with multiple criteria or objectives, alternative models of uncertainty such as fuzzy optimization, new insights into sensitivity analysis and application-specific results on topics that remain of prime relevance today such as job-shop scheduling and portfolio management.

References

1. Alem, D.J., Morabito, R.: Production planning in furniture settings via robust optimization. Comput. Oper. Res. **39**(2), 139–150 (2012)
2. Alumur, S.A., Nickel, S., Saldanha-da Gama, F.: Hub location under uncertainty. Transp. Res. Part B Methodol. **46**(4), 529–543 (2012).
3. Alvarez, P.P., Vera, J.R.: Application of robust optimization to the sawmill planning problem. Ann. Oper. Res. **219**(1), 457–475 (2014)
4. Alvarez-Miranda, E., Ljubic, I., Toth, P.: A note on the Bertsimas & Sim algorithm for robust combinatorial optimization problems. 4OR-A Q. J. Oper. Res. **11**(4), 349–360 (2013)
5. Ang, M., Lim, Y.F., Sim, M.: Robust storage assignment in unit-load warehouses. Manag. Sci. **58**(11), 2114–2130 (2012)
6. Aouam, T., Brahimi, N.: Integrated production planning and order acceptance under uncertainty: a robust optimization approach. Eur. J. Oper. Res. **228**(3), 504–515 (2013)
7. Ashayeri, J., Ma, N., Sotirov, R.: Supply chain downsizing under bankruptcy: a robust optimization approach. Int. J. Prod. Econ. **154**, 1–15 (2014)
8. Bandi, C., Bertsimas, D.: Tractable stochastic analysis in high dimensions via robust optimization. Math. Program. **134**(1, SI), 23–70 (2012)
9. Bandi, C., Bertsimas, D.: Optimal design for multi-item auctions: a robust optimization approach. Math. Oper. Res. **39**(4), 1012–1038 (2014)
10. Baron, O., Milner, J., Naseraldin, H.: Facility location: a robust optimization approach. Prod. Oper. Manag. **20**(5), 772–785 (2011)
11. Bayraksan, G., Love, D.: Data-driven stochastic programming using phi-divergences. INFORMS TutORials on operations research, pp. 1–19 (2015)
12. Beliakov, G., Kelarev, A.: Global non-smooth optimization in robust multivariate regression. Optim. Methods Softw. **28**(1), 124–138 (2013)
13. Ben-Tal, A., den Hertog, D.: Immunizing conic quadratic optimization problems against implementation errors. Technical Report. Tilburg University (2011)
14. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. Math. Oper. Res. **23**, 769–805 (1998)
15. Ben-Tal, A., Nemirovski, A.: Robust solutions of uncertain linear programs. Oper. Res. Lett. **25**, 1–13 (1999)
16. Ben-Tal, A., Nemirovski, A.: Robust solutions of linear programming problems contaminated with uncertain data. Math. Program. **88**, 411–424 (2000)
17. Ben-Tal, A., Nemirovski, A.: Selected topics in robust convex optimization. Math. Program. **112**, 125–158 (2011)
18. Ben-Tal, A., El-Ghaoui, L., Nemirovski, A.: Robust optimization. Princeton Series in Applied Mathematics. Princeton University Press, Princeton (2009)
19. Ben-Tal, A., Bhadra, S., Bhattacharyya, C., Nath, J.S.: Chance constrained uncertain classification via robust optimization. Math. Program. **127**(1, SI), 145–173 (2011)
20. Ben-Tal, A., Do Chung, B., Mandala, S.R., Yao, T.: Robust optimization for emergency logistics planning: risk mitigation in humanitarian relief supply chains. Transp. Res. B Methodol. **45**(8, SI), 1177–1189 (2011)
21. Ben-Tal, A., den Hertog, D., De Waegenaere, A., Melenberg, B., Rennen, G.: Robust solutions of optimization problems affected by uncertain probabilities. Manag. Sci. **59**(2), 341–357 (2013)

22. Ben-Tal, A., den Hertog, D., Vial, J.P.: Deriving robust counterparts of nonlinear uncertain inequalities. Math. Program. **149**, 265–299 (2015)
23. Ben-Tal, A., Hazan, E., Koren, T., Mannor, S.: Oracle-based robust optimization via online learning. Oper. Res. **63**(3), 628–638 (2015)
24. Bertsimas, D., Goyal, V.: On the approximability of adjustable robust convex optimization under uncertainty. Math. Meth. Oper. Res. **77**(3), 323–343 (2013)
25. Bertsimas, D., Sim, M.: The price of robustness. Oper. Res. **52**(1), 35–53 (2004)
26. Bertsimas, D., Takeda, A.: Optimizing over coherent risk measures and non-convexities: a robust mixed integer optimization approach. Comput. Optim. Appl. **62**(3), 613–639 (2015)
27. Bertsimas, D., Goyal, V., Sun, X.: A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. Math. Oper. Res. **36**(1), 24–54 (2011)
28. Bertsimas, D., Goyal, V., Lu, B.Y.: A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. Math. Program. **150**(2), 281–319 (2015)
29. Cacchiani, V., Toth, P.: Nominal and robust train timetabling problems. Eur. J. Oper. Res. **219**(3), 727–737 (2012)
30. Cacchiani, V., Caprara, A., Fischetti, M.: A Lagrangian heuristic for robustness, with an application to train timetabling. Transp. Sci. **46**(1), 124–133 (2012)
31. Caramanis, C., Mannor, S., Xu, H.: Robust optimization in machine learning. In: Optimization in Machine Learning, pp. 369–402 MIT Press, Cambridge, MA (2011)
32. Carlsson, D., Flisberg, P., Roennqvist, M.: Using robust optimization for distribution and inventory planning for a large pulp producer. Comput. Oper. Res. **44**, 214–225 (2014)
33. Chan, T.C.Y., Misic, V.V.: Adaptive and robust radiation therapy optimization for lung cancer. Eur. J. Oper. Res. **231**(3), 745–756 (2013)
34. Chan, T.C.Y., Mahmoudzadeh, H., Purdie, T.G.: A robust-CVaR optimization approach with application to breast cancer therapy. Eur. J. Oper. Res. **238**(3), 876–885 (2014)
35. Charnes, A., Cooper, W.W.: Chance-constrained programming. Manag. Sci. **6**(1), 73–79 (1959)
36. Chassein, A.B., Goerigk, M.: A new bound for the midpoint solution in minmax regret optimization with an application to the robust shortest path problem. Eur. J. Oper. Res. **244**(3), 739–747 (2015)
37. Che, A., Feng, J., Chen, H., Chu, C.: Robust optimization for the cyclic hoist scheduling problem. Eur. J. Oper. Res. **240**(3), 627–636 (2015)
38. Chen, L., He, S., Zhang, S.: Tight bounds for some risk measures, with applications to robust portfolio selection. Oper. Res. **59**(4), 847–865 (2011)
39. Classen, G., Koster, A.M.C.A., Schmeink, A.: A robust optimisation model and cutting planes for the planning of energy-efficient wireless networks. Comput. Oper. Res. **40**(1), 80–90 (2013)
40. Delage, E., Iancu, D.: Robust multistage decision-making. INFORMS TutORials, pp. 20–46 (2015)
41. Doan, X.V., Kruk, S., Wolkowicz, H.: A robust algorithm for semidefinite programming. Optim. Methods Softw. **27**(4–5, SI), 667–693 (2012)
42. Dupacova, J., Kopa, M.: Robustness in stochastic programs with risk constraints. Ann. Oper. Res. **200**(1), 55–74 (2012)
43. Duzgun, R., Thiele, A.: Robust binary optimization using a safe tractable approximation. Oper. Res. Lett. **43**(4), 445–449 (2015)
44. Duzgun, R., Thiele, A.: Robust optimization with multiple ranges: theory and application to pharmaceutical project selection. In: Proceedings of the 14th INFORMS ICS Conference, pp. 103–118 (2015)
45. El-Ghaoui, L., Lebret, H.: Robust solutions to least-squares problems with uncertain data. SIAM J. Matrix Anal. Appl. **18**(4), 1035–1064 (1997)
46. El-Ghaoui, L., Oustry, F., Lebret, H.: Robust solutions to uncertain semidefinite programs. SIAM J. Optim. **9**(1), 33–52 (1998)

47. Fang, J., Gao, Y., Sun, G., Xu, C., Li, Q.: Multiobjective robust design optimization of fatigue life for a truck cab. Reliab. Eng. Syst. Saf. **135**, 1–8 (2015)
48. Fertis, A., Baes, M., Luethi, H.J.: Robust risk management. Eur. J. Oper. Res. **222**(3), 663–672 (2012)
49. Fliege, J., Werner, R.: Robust multiobjective optimization & applications in portfolio optimization. Eur. J. Oper. Res. **234**(2), 422–433 (2014)
50. Gabrel, V., Murat, C., Thiele, A.: Recent advances in robust optimization: an overview. Eur. J. Oper. Res. **235**(3), 471–483 (2014)
51. Gabrel, V., Lacroix, M., Murat, C., Remli, N.: Robust location transportation problems under uncertain demands. Discret. Appl. Math. **164**(1, SI), 100–111 (2014). 1st International Symposium on Combinatorial Optimization (ISCO), Hammamet, TUNISIA, Mar 24–26, 2010
52. Galli, L.: Combinatorial and robust optimisation models and algorithms for railway applications. 4OR-A Q. J. Oper. Res. **9**(2), 215–218 (2011)
53. Ghezavati, V., Soltanzadeh, F., Hafezalkotob, A.: Optimization of reliability for a hierarchical facility location problem under disaster relief situations by a chance-constrained programming and robust optimization. Proc. Inst. Mech. Eng. O J Risk Reliab. **229**(6), 542–555 (2015)
54. Goh, J., Sim, M.: Distributionally robust optimization and its tractable approximations. Oper. Res. **58**, 902–917 (2010)
55. Golias, M., Portal, I., Konur, D., Kaisar, E., Kolomvos, G.: Robust berth scheduling at marine container terminals via hierarchical optimization. Comput. Oper. Res. **41**, 412–422 (2014)
56. Gorissen, B.L., Yanikoglu, I., den Hertog, D.: A practical guide to robust optimization. OMEGA Int. J. Manag. Sci. **53**, 124–137 (2015)
57. Goryashko, A.P., Nemirovski, A.S.: Robust energy cost optimization of water distribution system with uncertain demand. Autom. Remote Control **75**(10), 1754–1769 (2014)
58. Guelpinar, N., Pachamanova, D., Canakoglu, E.: Robust strategies for facility location under uncertainty. Eur. J. Oper. Res. **225**(1), 21–35 (2013)
59. Hanasusanto, G., Kuhn, D., Wallace, S., Zymler, S.: Distributional robust multi-item newsvendor problems with multimodal demand distributions. Math. Prog. **152**, 1–32 (2015)
60. Hanasusanto, G., Kuhn, D., Wiesemann, W.: K-adaptability in two-stage robust binary programming. Oper. Res. **63**, 877–891 (2015)
61. Hassanzadeh, F., Nemati, H., Sun, M.: Robust optimization for interactive multiobjective programming with imprecise information applied to R&D project portfolio selection. Eur. J. Oper. Res. **238**(1), 41–53 (2014)
62. Holte, M., Mannino, C.: The implementor/adversary algorithm for the cyclic and robust scheduling problem in health-care. Eur. J. Oper. Res. **226**, 551–559 (2013)
63. Houska, B., Diehl, M.: Nonlinear robust optimization via sequential convex bilevel programming. Math. Prog. **142**(1–2), 539–577 (2013)
64. Hu, J., Mehrotra, S.: Robust and stochastically weighted multiobjective optimization models and reformulations. Oper. Res. **60**(4), 936–953 (2012)
65. Hu, J., Mehrotra, S.: Robust decision making over a set of random targets or risk-averse utilities with an application to portfolio optimization. IIE Trans. **47**(4, SI), 358–372 (2015)
66. Iancu, D.A., Trichakis, N.: Pareto efficiency in robust optimization. Manag. Sci. **60**(1), 130–147 (2014)
67. Iancu, D.A., Sharma, M., Sviridenko, M.: Supermodularity and affine policies in dynamic robust optimization. Oper. Res. **61**(4), 941–956 (2013)
68. Ide, J., Tiedemann, M., Westphal, S., Haiduk, F.: An application of deterministic and robust optimization in the wood cutting industry. 4OR-A Q. J. Oper. Res. **13**(1), 35–57 (2015)
69. Jeyakumar, V., Li, G.Y.: Trust-region problems with linear inequality constraints: exact SDP relaxation, global optimality and robust optimization. Math. Prog. **147**(1–2), 171–206 (2014)
70. Jeyakumar, V., Li, G., Suthaharan, S.: Support vector machine classifiers with uncertain knowledge sets via robust optimization. Optimization **63**(7), 1099–1116 (2014)
71. Jiang, H., Netessine, S., Savin, S.: Technical note: Robust newsvendor competition under asymmetric information. Oper. Res. **59**, 254–261 (2011)

72. Jiang, R., Wang, J., Guan, Y.: Robust unit commitment with wind power and pumped storage hydro. IEEE Trans. Power Syst. **27**(2), 800–810 (2012)
73. Kakouris, I., Rustem, B.: Robust portfolio optimization with copulas. Eur. J. Oper. Res. **235**(1), 28–37 (2014)
74. Kang, S.C., Brisimi, T.S., Paschalidis, I.C.: Distribution-dependent robust linear optimization with applications to inventory control. Ann. Oper. Res. **231**(1), 229–263 (2015)
75. Kapsos, M., Christofides, N., Rustem, B.: Worst-case robust omega ratio. Eur. J. Oper. Res. **234**, 499–507 (2014)
76. Kawas, B., Thiele, A.: A log-robust optimization approach to portfolio management. OR Spectr. **33**(1), 207–233 (2011)
77. Kawas, B., Thiele, A.: Short sales in Log-robust portfolio management. Eur. J. Oper. Res. **215**(3), 651–661 (2011)
78. Kawas, B., Laumanns, M., Pratsini, E.: A robust optimization approach to enhancing reliability in production planning under non-compliance risks. OR Spectr. **35**(4), 835–865 (2013)
79. Koebis, E.: On robust optimization relations between scalar robust optimization and unconstrained multicriteria optimization. J. Optim. Theory Appl. **167**(3), 969–984 (2015)
80. Kuhn, D., Wiesemann, W., Georghiou, A.: Primal and dual linear decision rules in stochastic and robust optimization. Math. Prog. **130**(1), 177–209 (2011)
81. Lagos, G., Espinoza, D., Moreno, E., Vielma, J.P.: Restricted risk measures and robust optimization. Eur. J. Oper. Res. **241**(3), 771–782 (2015)
82. Lassiter, K., Khademi, A., Taaffe, K.M.: A robust optimization approach to volunteer management in humanitarian crises. Int. J. Prod. Econ. **163**, 97–111 (2015)
83. Li, M., Gabriel, S.A., Shim, Y., Azarm, S.: Interval uncertainty-based robust optimization for convex and non-convex quadratic programs with applications in network infrastructure planning. Netw. Spat. Econ. **11**(1), 159–191 (2011)
84. Lima, R.M., Novais, A.Q., Conejo, A.J.: Weekly self-scheduling, forward contracting, and pool involvement for an electricity producer. An adaptive robust optimization approach. Eur. J. Oper. Res. **240**(2), 457–475 (2015)
85. Long, D.Z., Qi, J.: Distributionally robust discrete optimization with entropic value-at-risk. Oper. Res. Lett. **42**(8), 532–538 (2014)
86. Maillet, B., Tokpavi, S., Vaucher, B.: Global minimum variance portfolio optimisation under some model risk: a robust regression-based approach. Eur. J. Oper. Res. **244**(1), 289–299 (2015)
87. Marzban, S., Mahootchi, M., Khamseh, A.A.: Developing a multi-period robust optimization model considering American style options. Ann. Oper. Res. **233**(1), 305–320 (2015)
88. Meng, F., Qi, J., Zhang, M., Ang, J., Chu, S., Sim, M.: A robust optimization model for managing elective admission in a public hospital. Oper. Res. **63**, 1452–1467 (2015)
89. Minoux, M.: Two-stage robust optimization, state-space representable uncertainty and applications. RAIRO Oper. Res. **48**(4), 455–475 (2014)
90. Moon, Y., Yao, T.: A robust mean absolute deviation model for portfolio optimization. Comput. Oper. Res. **38**(9), 1251–1258 (2011)
91. Mulvey, J., Vanderbei, R., Zenios, S.: Robust optimization of large-scale systems. Oper. Res. **43**, 264–281 (1995)
92. Najafi, M., Eshghi, K., Dullaert, W.: A multi-objective robust optimization model for logistics planning in the earthquake response phase. Transp. Res. Part E Log. **49**(1), 217–249 (2013)
93. Nemirovski, A.: On safe tractable approximations of chance constraints. Eur. J. Oper. Res. **219**, 707–718 (2012)
94. Nguyen, T.D., Lo, A.W.: Robust ranking and portfolio optimization. Eur. J. Oper. Res. **221**(2), 407–416 (2012)
95. Nha, V.T., Shin, S., Jeong, S.H.: Lexicographical dynamic goal programming approach to a robust design optimization within the pharmaceutical environment. Eur. J. Oper. Res. **229**(2), 505–517 (2013)
96. Ozmen, A., Weber, G.W., Karimov, A.: A new robust optimization tool applied on financial data. Pac. J. Optim. **9**(3), 535–552 (2013)

97. Paul, J.A., Wang, X.J.: Robust optimization for United States Department of Agriculture food aid bid allocations. Transp. Res. Part E Log. **82**, 129–146 (2015)
98. Poss, M.: Robust combinatorial optimization with variable budgeted uncertainty. 4OR-A Q. J. Oper. Res. **11**(1), 75–92 (2013)
99. Qiu, R., Shang, J.: Robust optimisation for risk-averse multi-period inventory decision with partial demand distribution information. Int. J. Prod. Res. **52**(24), 7472–7495 (2014)
100. Qiu, R., Shang, J., Huang, X.: Robust inventory decision under distribution uncertainty: a CVaR-based optimization approach. Int. J. Prod. Econ. **153**, 13–23 (2014)
101. Remli, N., Rekik, M.: A robust winner determination problem for combinatorial transportation auctions under uncertain shipment volumes. Transp. Res. Part C Emerg. Technol. **35**(SI), 204–217 (2013)
102. Sauma, E., Traub, F., Vera, J.: A Robust optimization approach to assess the effect of delays in the connection-to-the-grid time of new generation power plants over transmission expansion planning. Ann. Oper. Res. **229**(1), 703–741 (2015)
103. Schoenlein, M., Makuschewitz, T., Wirth, F., Scholz-Reiter, B.: Measurement and optimization of robust stability of multiclass queueing networks: applications in dynamic supply chains. Eur. J. Oper. Res. **229**(1), 179–189 (2013)
104. Shapiro, A.: A dynamic programming approach to adjustable robust optimization. Oper. Res. Lett. **39**(2), 83–87 (2011)
105. Shishebori, D., Babadi, A.Y.: Robust and reliable medical services network design under uncertain environment and system disruptions. Transp. Res. Part E Log. **77**, 268–288 (2015)
106. Sniedovich, M.: Fooled by local robustness. Risk Anal. **32**, 1630–1637 (2012)
107. Soyster, A.L.: Convex programming with set-inclusive constraints and applications to inexact linear programming. Oper. Res. **21**(5), 1154–1157 (1973)
108. Sun, H., Gao, Z., Szeto, W.Y., Long, J., Zhao, F.: A distributionally robust joint chance constrained optimization model for the dynamic network design problem under demand uncertainty. Netw. Spat. Econ. **14**(3–4), 409–433 (2014)
109. Suzuki, S., Kuroiwa, D., Lee, G.M.: Surrogate duality for robust optimization. Eur. J. Oper. Res. **231**(2), 257–262 (2013)
110. Tajik, N., Tavakkoli-Moghaddam, R., Vahdani, B., Mousavi, S.M.: A robust optimization approach for pollution routing problem with pickup and delivery under uncertainty. J. Manuf. Syst. **33**(2), 277–286 (2014)
111. Tong, X., Wu, F.: Robust reward-risk ratio optimization with application in allocation of generation asset. Optimization **63**(11), 1761–1779 (2014)
112. Varas, M., Maturana, S., Pascual, R., Vargas, I., Vera, J.: Scheduling production for a sawmill: a robust optimization approach. Int. J. Prod. Econ. **150**, 37–51 (2014)
113. Wiesemann, W., Kuhn, D., Rustem, B.: Robust Markov decision processes. Math. Oper. Res. **38**(1), 153–183 (2013)
114. Wiesemann, W., Kuhn, D., Sim, M.: Distributionally robust convex optimization. Oper. Res. **62**(6), 1358–1376 (2014)
115. Xu, H., Caramanis, C., Mannor, S.: A distributional interpretation of robust optimization. Math. Oper. Res. **37**(1), 95–110 (2012)
116. Xu, H., Caramanis, C., Sanghavi, S.: Robust PCA via outlier pursuit. IEEE Trans. Inf. Theory **58**(5), 3047–3064 (2012)
117. Ye, K., Parpas, P., Rustem, B.: Robust portfolio optimization: a conic programming approach. Comput. Optim. Appl. **52**(2), 463–481 (2012)
118. Zeng, B., Zhao, L.: Solving two-stage robust optimization problems using a column-and-constraint generation method. Oper. Res. Lett. **41**(5), 457–461 (2013)
119. Zhen, L.: Task assignment under uncertainty: stochastic programming and robust optimisation approaches. Int. J. Prod. Res. **53**(5), 1487–1502 (2015)
120. Zugno, M., Conejo, A.J.: A robust optimization approach to energy and reserve dispatch in electricity markets. Eur. J. Oper. Res. **247**(2), 659–671 (2015)
121. Zymler, S., Kuhn, D., Rustem, B.: Distributionally robust joint chance constraints with second-order moment information. Math. Prog. **137**(1–2), 167–198 (2013)

Chapter 6
Robust Discrete Optimization Under Discrete and Interval Uncertainty: A Survey

Adam Kasperski and Paweł Zieliński

Abstract In this chapter a review of recent results on robust discrete optimization is presented. The most popular discrete and interval uncertainty representations are discussed. Various robust concepts are presented, namely the traditional minmax (regret) approach with some of its recent extensions, and several two-stage concepts. A special attention is paid to the computational properties of the robust problems considered.

6.1 Introduction

In this chapter we will be concerned with a class of discrete optimization problems defined as follows. We are given a finite set of elements $E = \{e_1, \ldots, e_n\}$ and a set of feasible solutions $\Phi \subseteq 2^E$. Each element $e_i \in E$ has a nonnegative cost c_i and we seek a feasible solution $X \in \Phi$ which minimizes the total cost $f(X) = \sum_{e_i \in X} c_i$. This traditional deterministic discrete optimization problem will be denoted by $\mathscr{P}$. The above formulation encompasses, for instance, an important class of network problems. Namely, E can be identified with the set of arcs of a network $G = (V, E)$ and Φ contains some objects in G such as $s-t$ paths, spanning trees, $s-t$ cuts, perfect matchings, or Hamiltonian cycles. We thus get the well known and

A. Kasperski
Faculty of Computer Science and Management, Wrocław University of Technology, Wrocław, Poland
e-mail: adam.kasperski@pwr.edu.pl

P. Zieliński (✉)
Faculty of Fundamental Problems of Technology, Wrocław University of Technology, Wrocław, Poland
e-mail: pawel.zielinski@pwr.edu.pl

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_6

basic problems such as SHORTEST PATH, MINIMUM SPANNING TREE, MINIMUM S-T CUT, MINIMUM ASSIGNMENT, or TRAVELING SALESPERSON, respectively. A comprehensive description of the class of deterministic network problems can be found, for example, in books [1, 78].

In most cases, $\mathscr{P}$ can be alternatively formulated as a 0-1 programming problem. Indeed, a binary variable $x_i \in \{0,1\}$ is associated with element $e_i \in E$ and so $\mathscr{P}$ has the following formulation:

$$\begin{array}{ll} \min & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} & (x_1,\ldots,x_n) \in ch(\Phi), \end{array}$$

where $ch(\Phi)$ is the set of characteristic vectors of Φ, described in a compact form by a system of constraints involving $x_1,\ldots,x_n$. For example, when we have one constraint of the form $\sum_{i=1}^{n} w_i x_i \geq p$, we obtain the KNAPSACK problem. If, additionally, $w_i = 1$ for each $i \in [n]$ and p is an integer in $[n]$ ($[n]$ denotes the set $\{1,\ldots,n\}$), then we get the SELECTION problem. An optimal solution to this problem can be computed in $O(n)$ time by choosing p elements out of E of the smallest costs. In this chapter, we will also discuss the following REPRESENTATIVES SELECTION problem (it is also called WEIGHTED HITTING DISJOINT SET, see, e.g., [17]). Let us partition the set $[n]$ into u disjoint subsets $T_1,\ldots,T_u$. Then $ch(\Phi)$ is described by a system of constraints of the form $\sum_{i\in T_j} x_i = 1$ for each $j \in [u]$. Hence, each feasible solution is composed of exactly one element e_j from each T_j. An important characteristic of this problem is the value of $r_{\max} = \max_{j\in[u]} |T_j|$. An optimal solution to this problem is composed of elements of the smallest costs from each T_j. Both SELECTION and REPRESENTATIVES SELECTION problems become nontrivial under uncertainty. We will discuss them later in detail as they allow us to obtain strong negative complexity results for many robust versions of discrete optimization problems.

In many practical applications the element costs are often uncertain, which means that their precise values are not known before computing a solution. In this case a *scenario set* $\mathscr{U}$, containing all possible realizations of the element costs, is a part of input. Each particular cost realization $(c_1^S,\ldots,c_n^S) \in \mathscr{U}$ is called a *scenario*. Then $f(X,S) = \sum_{e_i\in X} c_i^S$ is the cost of solution X under scenario S. In this chapter we will focus on two popular methods of defining set $\mathscr{U}$—*discrete and interval uncertainty representations*. For the *discrete uncertainty representation* [61], scenario set, denoted by $\mathscr{U}_D$, contains K explicitly listed scenarios. This uncertainty representation is appropriate when each scenario corresponds to an event which globally influences the element costs. For example, an uncertain weather forecast can globally change a system environment, and these uncertain weather conditions can be modeled by discrete scenarios. For the *interval uncertainty representation* [15], scenario set, denoted by $\mathscr{U}_I^\ell$ is defined as follows. We assume that the cost of element e_i can take any value within the interval $[c_i, c_i + d_i]$, where c_i is a nominal cost and d_i is the maximum deviation of the value of the cost from its nominal value. Then $\mathscr{U}_I^\ell$ is a subset of the Cartesian product of these intervals, under the additional assumption that in each scenario in $\mathscr{U}_I^\ell$, the costs of at most ℓ elements can be greater than their nominal values. The value of $\ell \in [0,n]$ is fixed and allows us to control

the degree of uncertainty. When $\ell = 0$, then we get a deterministic problem with one scenario. On the other hand, when $\ell = n$, then we get the traditional interval uncertainty representation [61], in which scenario set is equal to the Cartesian product of all the uncertainty intervals. We will denote this particular special case of scenario set by $\mathscr{U}_I = \mathscr{U}_I^n$. The scenario set $\mathscr{U}_I$ models a local uncertainty, i.e. we assume that the cost of each element may vary independently on the costs of the remaining elements. For instance, a traveling time of some link is often uncertain and can be modeled by a closed interval which provides us a bound on the minimum and the maximum possible value of the traveling time. It is often not possible to measure some costs precisely and the measurement error can also be expressed as a closed interval.

In mathematical programming problems some other types of scenario sets, in particular the *ellipsoidal uncertainty* or the *column-wise uncertainty set*, are also used. In general $\mathscr{U}$ can be any set, typically assumed to be convex [14]. In this chapter we will not be concerned with such more general scenario sets. Some discussion on them can be found in the recent survey [38]. In robust optimization, also the set of feasible solutions can be uncertain and may depend on a scenario (see, e.g. [65]). In the class of problems discussed in this chapter the set of feasible solutions Φ is deterministic, i.e. it remains the same for each scenario in $\mathscr{U}$. Under this assumption, the discrete and interval uncertainty representations are the easiest and, in many cases, possess sufficient expressive power.

If no additional information for $\mathscr{U}$ (such as a probability distribution) is provided, then we face a decision problem under *uncertainty*. In order to choose a solution we can use some well known criteria used in decision theory under uncertainty (see e.g. [63]). Among them there are the minmax and minmax regret criteria, which assume that the decision maker is risk averse and seeks a solution minimizing the cost or opportunity loss in a worst case, i.e under a worst scenario which may occur. By using the minmax (regret) criterion we obtain the *robust minmax (regret) optimization problem*. This traditional robust approach to discrete optimization has some well known drawbacks, which we will discuss in more detail in Sect. 6.2. By applying the minmax (regret) criterion we may sometimes get unreasonable solutions (we will show some examples in Sect. 6.2). Furthermore, it is not true that decision makers are always extremely risk averse. Hence, there is a need to soften the very conservative minmax (regret) criterion. Also, in many practical applications decision makers have some additional information provided with $\mathscr{U}$. For example, a probability distribution in $\mathscr{U}$ or its estimation may be available. This information should be taken into account while computing a solution. In Sect. 6.3 we will present some recent extensions of the robust approach which take into account both an attitude of decision makers towards a risk and an information about the probability distribution in $\mathscr{U}$.

The minmax approach can be generalized by considering the *robust optimization problem with incremental recourse* [72]. This problem can be seen as a zero-sum game against the nature with the following rules. The decision maker chooses first a solution X whose cost $f(X)$ is precisely known. Then nature picks a scenario S from $\mathscr{U}$ and the decision maker, chooses the next solution Y after observing S.

The solution Y has the cost $f(Y,S)$ and must be of some predefined distance from X. The decision maker wants to minimize the total cost $f(X)+f(Y,S)$ while the nature aims to maximize this total cost, i.e. it always picks the worst scenario for solution X. It is easily seen that the robust optimization problem with incremental recourse contains the minmax problem as a special case. Indeed, by assuming that the initial cost of X is always 0 and Y must be the same as X (no modification of X is allowed) we arrive to the minmax problem. The robust optimization with incremental recourse is similar to *robust recoverable* optimization [17, 18, 62], because a limited recovery action is allowed after observing which scenario has occurred. We will study the robust optimization problems with incremental recourse in Sect. 6.4.

The traditional min-max (regret) approach is a one-stage decision problem, i.e. a complete solution must be computed before a true scenario reveals. However, many practical problems have a two-stage nature. Namely, a partial solution is formed in the first stage, when the costs are precisely known and then it is completed optimally when a true cost scenario from $\mathscr{U}$ occurs. We seek a solution whose maximum total cost in both stages is minimum. We will discuss the class of robust two-stage problems in Sect. 6.5.

The aim of this chapter is to present and compare various concepts used in robust discrete optimization under the discrete and interval uncertainty representations. A survey of the results in the area of robust minmax (regret) optimization up to 2009 can be found in [5, 40, 61]. In this chapter we present new results and concepts which have recently appeared in the literature. We will pay a special attention to the computational properties of the problems under study (a recent survey from the algorithmic perspective can be found in [38]). In Sect. 6.2 we present the traditional minmax (regret) approach. We also show, in Sect. 6.3, some of its extensions which allow decision makers to model their attitude towards risk and exploit scenario probabilities. In Sect. 6.4 we examine the robust optimization problems with incremental recourse. Finally, in Sect. 6.5 we describe the class of robust two-stage problems.

The class of problems considered in this chapter is rather broad. However, it does not cover an important class of sequencing problems in which a feasible solution is represented by a permutation of the elements (typically called jobs). A recent survey of the results for the minmax (regret) sequencing problems can be found in [50]. Another class of problems, which is not discussed in detail, contains the ones with the bottleneck cost function. The minmax (regret) versions of such problems were investigated in [8], where it was shown that their complexity is nearly the same as the complexity of their deterministic counterparts. An extension of the minmax bottleneck problems has been discussed in [49]. We also do not mention about the maximum relative regret criterion. Some properties of this criterion, in particular its connections with the maximum regret, can be found in [10, 61].

6.2 Robust Min-Max (Regret) Problems

In this section we discus the traditional robust approach to deal with discrete optimization problems with uncertain costs. We describe the minmax and minmax regret criteria, which are typically used in the robust optimization framework. We present the known complexity results for basic problems and show some drawbacks of the minmax (regret) approach.

6.2.1 Using the Minmax Criterion

This section is devoted to the study of the following *minmax problem*:

$$\text{MIN-MAX } \mathscr{P}: \min_{X\in\Phi}\max_{S\in\mathscr{U}} f(X,S).$$

We thus seek a solution minimizing the maximum cost over all scenarios. Minmax is the most popular criterion used in robust optimization [14, 61]. The minmax problem can be alternatively stated as follows:

$$\begin{array}{ll} \min\ t & \\ \text{s.t. } \sum_{i\in[n]} c_i^S x_i \le t & \forall S\in\mathscr{U} \\ (x_1,\ldots,x_n)\in ch(\Phi) & \\ t\ge 0 & \end{array} \tag{6.1}$$

The minmax criterion can be extremely conservative and it will be used by pessimistic decision makers, or in situations in which it is very important to avoid bad scenarios. Perhaps, the most serious drawback of the minmax approach is that it may lead to solutions which are not Pareto optimal. Consider two sample MIN-MAX SHORTEST PATH problems, shown in Fig. 6.1. All the three paths, depicted in Fig. 6.1a, have the same maximum cost equal to 16. Hence, we can choose the path $\{e_2,e_5\}$ which is weakly dominated by the remaining two paths. When the number of scenarios becomes large, then the so-called *drowning effect* may occur [32], i.e. only one bad scenario is taken into account while choosing a solution and the information associated with the remaining scenarios is ignored. A similar situation occurs for the interval uncertainty representation and it is shown in the sample problem in Fig. 6.1b. Path $\{e_1\}$ is almost always better than path $\{e_2\}$, but both can be chosen after applying the minmax criterion. Also, an optimal minmax solution which is Pareto optimal, can be a questionable choice. Consider again the sample problem presented in Fig. 6.1a, and change the cost of arc e_2 under S_3 to 7. The path $\{e_2,e_5\}$ is then an optimal minmax solution which is also Pareto optimal. However, this path is only slightly better than $\{e_1,e_4\}$ under S_3 and much worse under S_1 and S_2.

The examples given in Fig. 6.1 show that there is a need of modification of the minmax criterion. If the decision maker is interested in minimizing the total cost,

then a chosen solution should always be Pareto optimal. Furthermore, an attitude of decision makers towards a risk should be taken into account, because not all decision makers are extremely risk averse. In Sect. 6.3.1 we will suggest a criterion which allows us to overcome both these drawbacks. In the next section, we will discuss all the known complexity results for MIN-MAX $\mathscr{P}$.

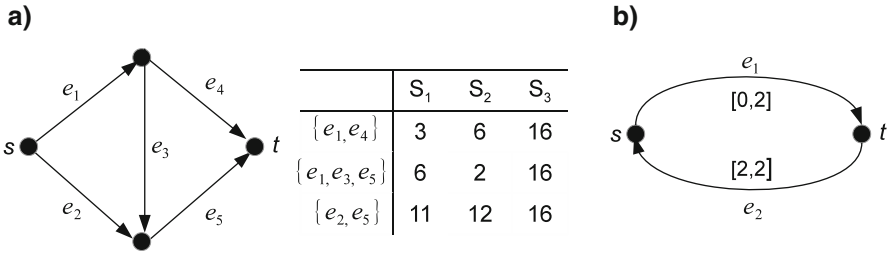
6.2.1.1 Discrete Uncertainty Representation

Consider the discrete uncertainty representation, i.e. when $\mathscr{U} = \mathscr{U}_D$. The known complexity results for some basic minmax problems are shown in Table 6.1.

Table 6.1: Complexity results for various MIN-MAX $\mathscr{P}$ problems with scenario set $\mathscr{U}_D$

MIN-MAX $\mathscr{P}$	Constant K	Unbounded K
SHORTEST PATH	NP-hard for $K = 2$ [89], FPTAS [3]	Strongly NP-hard [89], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [45], appr. within K [5]
MINIMUM SPANNING TREE	NP-hard for $K = 2$ [61, 88], FPTAS [3]	Strongly NP-hard [61, 88], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [48], appr. within $O(\log^2 n)$ with high probability [48]
MINIMUM S-T CUT	Strongly NP-hard for $K = 2$ [4]	Strongly NP-hard [4], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [45], appr. within K [5]
MINIMUM ASSIGNMENT	Strongly NP-hard for $K = 2$ [84, 90]	Strongly NP-hard [2, 84, 90], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [45], appr. within K [5]
SELECTION	NP-hard for $K = 2$ [9], FPTAS [3]	Strongly NP-hard [46], not appr. within any const. $\gamma > 0$ [57], appr. within $O(\log K / \log\log K)$ [30]
REPR. SELECTION	NP-hard for $K = 2$ [31], FPTAS [31]	Strongly NP-hard [31], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [59], not appr. within $2-\varepsilon$ when $r_{\max} = 2$ for any $\varepsilon > 0$ [28], appr. within $\min\{K, r_{\max}\}$ [59]
KNAPSACK	NP-hard for $K = 1$ [37], FPTAS [3]	Strongly NP-hard [46, 87], not appr. within any const. $\gamma > 0$ [57]

Observe that all these problems become NP-hard or strongly NP-hard, even when the number of scenarios equals 2. However, if the number of scenarios is constant then some of them can be solved in pseudopolynomial time (typically a dynamic programming method is applied) and admit a fully polynomial time approximation scheme (FPTAS). We should point out, here, that the running times of the



	S_1	S_2	S_3
$\{e_1, e_4\}$	3	6	16
$\{e_1, e_3, e_5\}$	6	2	16
$\{e_2, e_5\}$	11	12	16

Fig. 6.1: (**a**) A sample MIN-MAX SHORTEST PATH problem with three scenarios $S_1 = (2,10,3,1,1)$, $S_2 = (1,11,0,5,1)$, $S_3 = (8,8,0,8,8)$. (**b**) A sample MIN-MAX SHORTEST PATH problem with interval costs

pseudopolynomial algorithms and the FPTAS's proposed in the literature are exponential in K and, in consequence, the practical applicability of them is rather limited. The complexity of the problems become worse when the number of scenarios is a part of input. In particular, the network problems are then hard to approximate within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [45, 48]. A similar result holds for the MIN-MAX REPRESENTATIVES SELECTION problem [59]. The MIN-MAX SELECTION and MIN-MAX KNAPSACK problems are then hard to approximate within any constant factor $\gamma > 0$ [57].

If the underlying deterministic problem $\mathscr{P}$ is polynomially solvable, then MIN-MAX $\mathscr{P}$ is approximable within K. It is enough to solve the deterministic problem $\mathscr{P}$ for the aggregated costs $\hat{c}_i = \max_{S \in \mathscr{U}_D} c_i^S$ (or $\hat{c}_i = \sum_{S \in \mathscr{U}_D} c_i^S$), $i \in [n]$. A straightforward proof of this fact can be found, for instance, in [5]. This approximation ratio has been improved for two particular problems. For MIN-MAX MINIMUM SPANNING TREE a randomized $O(\log^2 n)$-approximation algorithm was constructed in [48] and for the MIN-MAX SELECTION problem a deterministic $O(\log K / \log\log K)$-approximation algorithm was proposed in [30]. These algorithms are based on the idea of randomized rounding of linear programming programs, which seems to be a promising tool for establishing stronger approximation results for the minmax problems, when the number of scenarios is a part of input.

The MIN-MAX $\mathscr{P}$ problem can be solved exactly by applying the formulation (6.1). After replacing $(x_1, \ldots, x_n) \in ch(\Phi)$ with a system of linear constraints, we obtain a compact MIP formulation for the problem. Other exact methods for this problem, such as branch and bound algorithms, can be found in [61].

In some cases, the underlying deterministic problem $\mathscr{P}$ is a maximization problem, i.e. we seek a solution which maximizes the total cost. It is then natural to study the symmetric MAX-MIN $\mathscr{P}$ problem, in which we wish to find a solution maximizing the minimum cost over all scenarios, i.e. $\max_{X \in \Phi} \min_{S \in \mathscr{U}} f(X,S)$. Interestingly, for scenario set $\mathscr{U}_D$, MAX-MIN $\mathscr{P}$ seems to be harder than the corresponding MIN-MAX $\mathscr{P}$ problem. In [52] it has been shown that MAX-MIN INDEPENDENT SET problem in interval graphs (this problem was first discussed in [75]), whose deterministic version is polynomially solvable, is not at all approximable when K is a part of input. A similar fact was observed for MAX-MIN KNAPSACK in [3] (see also [74]).

6.2.1.2 Interval Uncertainty Representation

Let us address the interval uncertainty representation, i.e. when $\mathscr{U} = \mathscr{U}_I^{\ell}$. We first discuss the case $\mathscr{U}_I = \mathscr{U}_I^n$. It is easy to check that the complexity of MIN-MAX $\mathscr{P}$ is then almost the same as $\mathscr{P}$, because it is sufficient to solve the deterministic problem $\mathscr{P}$ for scenario $(c_1 + d_1, \ldots, c_n + d_n)$. Consequently, when $\mathscr{P}$ is solvable in $O(T(n))$ time, then MIN-MAX $\mathscr{P}$ is solvable in $O(n + T(n))$ time. The problem is more challenging when $\mathscr{U} = \mathscr{U}_I^{\ell}$ for a fixed $\ell \in [0,n]$. An algorithm for this case was proposed in [15]. We now briefly describe it. Let us number the elements so that $d_1 \geq d_2 \geq \cdots \geq d_n$ and define $d_{n+1} = 0$. Define scenario S^j under which the cost of e_i is equal to $c_i + (d_i - d_j)$ if $i \leq j$ and c_i otherwise, where $j \in [n+1]$. In [15] it has been shown that MIN-MAX $\mathscr{P}$ with scenario set $\mathscr{U}_I^{\ell}$ is equivalent to the following problem:

$$\min_{j \in [n+1]} (\ell d_j + \min_{X \in \Phi} f(X, S^j)). \tag{6.2}$$

Observe that (6.2) reduces to solving $n+1$ deterministic problems $\mathscr{P}$ for the costs specified in scenarios $S^1, \ldots, S^{n+1}$ and, in consequence, when $\mathscr{P}$ is solvable in $O(T(n))$ time, then MIN-MAX $\mathscr{P}$ is solvable in $O(nT(n))$ time. We thus get a tractable class of problems under uncertainty. Furthermore, it has been observed in [15] that this algorithm can be extended to problems $\mathscr{P}$ which are NP-hard but admit an α-approximation algorithm. In this case, that approximation algorithm can be used to solve the inner problem $\min_{X \in \Phi} f(X, S^j)$ and the minmax problem is also approximable within α.

6.2.2 Using the Minmax Regret Criterion

In this section we treat the following *minmax regret problem*:

$$\text{MIN-MAX REGRET } \mathscr{P}: \min_{X \in \Phi} \max_{S \in \mathscr{U}} (f(X,S) - f^*(S)),$$

where $f^*(S)$ is the cost of an optimal solution under scenario S. The quantity $f(X,S) - f^*(S)$ is called a *regret* of X under S and it expresses a deviation of solution X from the optimum under S. We thus seek a solution which minimizes the maximum regret over all scenarios. The maximum regret criterion is also called *Savage criterion* or *maximum opportunity loss*. The minmax regret problem can be alternatively stated as follows:

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \sum_{i \in [n]} c_i^S x_i \leq t + t^S \quad \forall S \in \mathscr{U} \\ & (x_1, \ldots, x_n) \in ch(\Phi) \\ & t \geq 0, \end{array} \tag{6.3}$$

where t^S is the cost of an optimal solution under scenario S. If we apply the minmax regret criterion to the sample problem presented in Fig. 6.1, then we get the reasonable paths $\{e_1, e_3, e_5\}$ in Fig. 6.1a and $\{e_1\}$ in Fig. 6.1b. Observe that the maximum regret of $\{e_1\}$ equals 0 which means that this path is optimal under each scenario.

It is important to realize that the maximum regret is quite different quantity than the maximum cost. In the former, the decision maker aims to minimize the opportunity loss, i.e. the cost of a solution is compared ex-post to the cost of the best solution which could be chosen. Consider the sample MIN-MAX REGRET SHORTEST PATH problem depicted in Fig. 6.2a. Both paths $\{e_1\}$ and $\{e_2\}$ have the same maximum regret equal to 1. However, the maximum cost of path $\{e_1\}$ is twice the maximum cost of $\{e_2\}$. Hence, a solution with small maximum regret may have a large maximum cost in comparison with other solutions. Decision makers who just want to minimize the solution cost should be careful while using the minmax regret criterion.

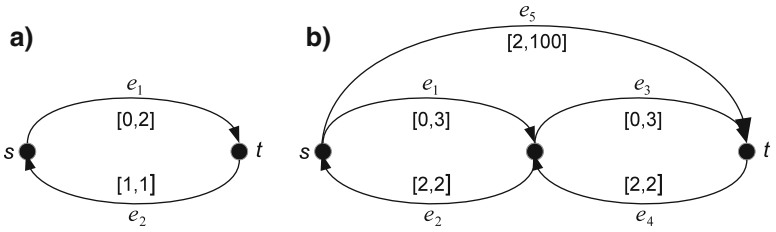


Fig. 6.2: Two sample MIN-MAX REGRET SHORTEST PATH problems with scenarios sets $\mathscr{U}_I$. (a) An instance with two paths that have the same maximum regret and different maximum costs. (b) An instance for which the maximum regret criterion does not satisfy the property of independency of irrelevant alternatives.

Another drawback of the minmax regret criterion is shown in Fig. 6.2b. It is easy to check that path $\{e_1, e_4\}$ has the smallest maximum regret equal to 3 and there is no path with smaller maximum regret. Suppose that we remove the path (a single arc) $\{e_5\}$ from the network. Then path $\{e_1, e_3\}$ has the smallest maximum regret equal to 2 and no other path has smaller maximum regret. Observe that the path $\{e_5\}$ is never optimal since its regret is very large. This example shows that the maximum regret criterion does not satisfy the property of *independency of irrelevant alternatives* [63], i.e. adding a non-optimal (and thus irrelevant) solution to the problem can make the optimal solution nonoptimal and vice versa.

It is evident that MIN-MAX REGRET $\mathscr{P}$ is NP-hard and not at all approximable when $\mathscr{P}$ is NP-hard. This is true even in the deterministic case when $K = 1$. It follows from the fact that it is then NP-hard to compute a solution of the maximum regret equal to 0. Also, computing the maximum regret of a given solution is, in this case, NP-hard. This implies, in particular, that MIN-MAX REGRET KNAPSACK is not at all approximable under both discrete and interval uncertainty representations.

In [64] the following randomized version of MIN-MAX REGRET $\mathscr{P}$ has been proposed. That is, instead of choosing a single solution, a probability distribution over all solutions is computed and we seek a probability distribution which mini-

mizes the maximum expected regret. This problem can be seen as a game, in which the decision maker chooses a probability distribution and an adversary chooses then the worst scenario, knowing this probability distribution. Notice that in the traditional minmax regret problem, the decision maker is restricted to choose the solution deterministically. Interestingly, the best probability distribution can be computed in polynomial time if $\mathscr{P}$ is polynomially solvable. This holds for both discrete and interval uncertainty representations (see [64] for details).

6.2.2.1 Discrete Uncertainty Representation

Let us now discuss the discrete uncertainty representation, i.e. the case when $\mathscr{U} = \mathscr{U}_D$. The known complexity results for the minmax regret versions of some basic problems $\mathscr{P}$ are shown in Table 6.2.

Table 6.2: Complexity results for various MIN-MAX REGRET $\mathscr{P}$ problems with scenario set $\mathscr{U}_D$

MIN-MAX REGRET $\mathscr{P}$	Constant K	Unbounded K
SHORTEST PATH	NP-hard for $K = 2$ [89], FPTAS [3]	Strongly NP-hard [89], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [45], appr. within K [5]
MINIMUM SPANNING TREE	NP-hard for $K = 2$ [61, 88], FPTAS [3]	Strongly NP-hard [61, 88], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [48], appr. within K [5]
MINIMUM S-T CUT	Strongly NP-hard for $K = 2$ [4]	Strongly NP-hard [4], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [45], appr. within K [5]
MINIMUM ASSIGNMENT	Strongly NP-hard for $K = 2$ [84, 90]	Strongly NP-hard [2, 84, 90], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [45], appr. within K [5]
SELECTION	NP-hard for $K = 2$ [9], FPTAS [3]	Strongly NP-hard [46], not appr. within any const. $\gamma > 0$ [57], appr. within K [5]
REPR. SELECTION	NP-hard for $K = 2$ [31], FPTAS [31]	Strongly NP-hard [31], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [59], not appr. within $2-\varepsilon$ when $r_{\max} = 2$ for any $\varepsilon > 0$ [28], appr. within K [5]
KNAPSACK	NP-hard for $K = 1$ [37], not at all appr.	Strongly NP-hard [46, 87], not at all appr.

Similarly to MIN-MAX $\mathscr{P}$, MIN-MAX REGRET $\mathscr{P}$ becomes NP-hard or strongly NP-hard when the number of scenarios equals 2. These negative results can be

strengthen when the number of scenarios is a part of input and they are the same as for MIN-MAX $\mathscr{P}$ (see Table 6.1). In fact, the proof showing the hardness of MIN-MAX $\mathscr{P}$ can be, in most cases, easily modified to show the same hardness result for the minmax regret version of $\mathscr{P}$. Typically, it suffices to add a number of dummy elements and scenarios to the constructed instance.

Observe that there is lack of stronger positive results when the number of scenarios is a part of input. The only known and general result states that when $\mathscr{P}$ is polynomially solvable, then MIN-MAX REGRET $\mathscr{P}$ is approximable within K. The idea is to solve the deterministic problem for the average costs $\hat{c}_i = \frac{1}{K}\sum_{S\in\mathscr{U}_D} c_i^S$, $i \in [n]$. A straightforward proof of this fact can be found in [5].

We can use (6.3) to construct a compact MIP formulation for the minmax regret problem. However, the underlying problem $\mathscr{P}$ must by polynomially solvable, since we need the costs of the optimal solutions t^S for each $S \in \mathscr{U}_D$. Other exact methods for solving the problem can be found in [61].

6.2.2.2 Interval Uncertainty Representation

We now turn to case when $\mathscr{U} = \mathscr{U}_I$, i.e. the interval uncertainty representation. In the existing literature the problem with scenario set $\mathscr{U}_I$ has been extensively studied. To the best of our knowledge, more general scenario set $\mathscr{U}_I^\ell$ has been not yet investigated. The known complexity results and solution methods for various problems, under scenario set $\mathscr{U}_I$, are shown in Table 6.3.

The number of scenarios in $\mathscr{U}_I$ is infinite. It is, however, easy to show that we can replace $\mathscr{U}_I$ with the set of *extreme scenarios*, which is the Cartesian product $\Pi_{i\in[n]}\{c_i, c_i + d_i\}$. It is also not difficult to show (see, e.g. [40, 47]) that the maximum regret of X equals $f(X, S_X) - f^*(S_X)$, where S_X is the extreme scenario under which the costs of $e_i \in X$ equal $c_i + d_i$ and the costs of $e_i \notin X$ are equal to c_i. Consequently, the maximum regret of a given solution X can be computed in polynomial time if the underlying deterministic problem $\mathscr{P}$ is polynomially solvable. Remarkable, this is not the case for the minmax regret version of the linear programming problem with interval objective function coefficients, since it has been shown in [12] that computing the maximum regret of a given solution is strongly NP-hard.

It turns out (see [11]) that in order to compute an optimal minmax regret solution it is enough to compute an optimal solution for each extreme scenario and choose the best one. Consequently, if the number of *nondegenerate* cost intervals, i.e. such that $d_i > 0$, is bounded by $r \cdot \log n$, then it is sufficient to enumerate at most n^r solutions. This yields a polynomial method for constant r. Obviously, this method is exponential in general case.

Let us now discuss some general properties of MIN-MAX REGRET $\mathscr{P}$. A solution X is called *possibly optimal* if it is optimal under at least one scenario in $\mathscr{U}$ and X is called *necessarily optimal* if it is optimal under all scenarios in $\mathscr{U}$. Similarly, an element e_i is *possibly optimal* if it is a part of an optimal solution under at least one scenario and it is *necessarily optimal* if it is a part of an optimal solution under all scenarios. It turns out that under scenario set $\mathscr{U}_I$, each optimal minmax regret

Table 6.3: Complexity results and solutions methods for various MIN-MAX REGRET $\mathscr{P}$ problems with scenario set $\mathscr{U}_I$

MIN-MAX REGRET $\mathscr{P}$	Complexity	Solution methods
SHORTEST PATH	Strongly NP-hard [11], appr. within 2 [41], NP-hard for planar graphs [91], NP-hard for sp-graphs [42], FPTAS for sp-graphs [44]	MIP [39], B&B [23, 70], Benders [69], Enumeration [67], Other methods [33]
MINIMUM SPANNING TREE	Strongly NP-hard [6, 11], appr. within 2 [41]	MIP [86], B&B [7, 68], B&Cut [81], Benders [66], Tabu Search [56], Simulated Annealing [73], Other methods [33]
MINIMUM S-T CUT	Strongly NP-hard [4], appr. within 2 [41], NP-hard for sp-graphs [42, 44], FPTAS for sp-graphs [44]	MIP [40]
MINIMUM ASSIGNMENT	Strongly NP-hard [2], appr. within 2 [41]	MIP [40], Benders [79], Local Search [79], GA [79]
SELECTION	Solv. in $O(n \cdot \min\{n, n-p\})$ time [24]	
REPR. SELECTION	Solv. in $O(n^2)$ time [31]	
KNAPSACK	Σ_2^p-hard [27], not at all appr.	MIP [36], B&Cut [36], Local Search [36]

solution is possibly optimal and is entirely composed of possibly optimal elements. This fact was first observed for the minmax regret versions of SHORTEST PATH and MINIMUM SPANNING TREE in [39, 86] and it was generalized to all problems $\mathscr{P}$ in [47]. Notice, that this is not the case for scenario set $\mathscr{U}_D$, where it is easy to construct a sample problem whose optimal minmax regret solution is not optimal under any scenario. On the other hand, the maximum regret of a necessarily optimal solution equals 0, so it must be the optimal minmax regret solution (this is true for any scenario set $\mathscr{U}$). It was shown in [47] that when all cost intervals are nondegenerate, i.e. $d_i > 0$ for all $i \in [n]$, then there is an optimal minmax regret solution containing all necessarily optimal elements.

The notions of possibly and necessarily optimal elements can be very useful, as they allow us to reduce the size of a problem instance before a solution is computed, for example, by using a MIP formulation or a branch and bound algorithm. Namely, all non-possibly optimal elements can be removed from E and, under the absence of degeneracy, all necessarily elements can be automatically added to the solution constructed. Some computational tests (see. e.g. [39, 56]) suggest that for many instances more than 50 % elements are non-possibly optimal. One can also expect several elements to be necessarily optimal in each instance. Hence, a partial solution can be formed before a more complex algorithm is executed. Unfortunately, detecting possibly and necessarily optimal elements is not an easy task in general. In particular, the problem of checking whether a given element is possibly optimal, is strongly NP-hard for the SHORTEST PATH, MINIMUM ASSIGNMENT and MIN-

IMUM S-T CUT problems [21, 47]. All possibly and necessarily optimal elements can be detected in polynomial time if $\mathscr{P}$ is a matroidal problem [43], in particular, when $\mathscr{P}$ is SELECTION or MINIMUM SPANNING TREE. For the SHORTEST PATH problem a subset of possibly optimal elements (arcs) can be detected by efficient algorithms proposed in [20, 39]. Also, when the network is acyclic all necessarily optimal arcs can be detected in polynomial time [35]

As we can see in Table 6.3, the minmax regret versions of all the basic network problems are strongly NP-hard for general graphs. Two special cases, namely MIN-MAX REGRET SHORTEST PATH and MIN-MAX REGRET MINIMUM S-T CUT in series-parallel multidigraphs can be solved in pseudopolynomial time and admit an FPTAS [42, 44]. Fortunately, the following positive and general result is known for all polynomially solvable problems $\mathscr{P}$. Let S^M be the *midpoint scenario*, under which the cost of e_i is equal to $c_i + 0.5d_i$, $i \in [n]$. In [41] it has been shown that if X^* is an optimal solution under scenario S^M, then the maximum regret of X^* is at most twice the maximum regret of an optimal minmax regret solution. Consequently, if $\mathscr{P}$ is polynomially solvable, then MIN-MAX REGRET $\mathscr{P}$ is approximable within 2. The 2-approximation algorithm has been extended to a wider class of minmax regret problems in [25]. We do not know whether there exists an approximation algorithm with a performance ratio better than 2 (except for some very special cases). Also, no negative approximation result for MIN-MAX REGRET $\mathscr{P}$ is known, when $\mathscr{P}$ is polynomially solvable. So, the existence of an PTAS in this case cannot be excluded. Observe that this result allows us to detect efficiently a solution with the maximum regret equal to 0 (i.e. a necessarily optimal solution).

The computational tests (see, e.g. [40, 56]) suggest that the approximation algorithm behaves well in practice. It is often profitable to modify it by considering two solutions: an optimal solution under the midpoint scenario and an optimal solution under the pessimistic scenario $(c_1 + d_1, \ldots, c_n + d_n)$. The approximation algorithm, denoted as AMU, returns the better of these two solutions. Algorithm AMU seems to perform well except for some rather artificial instances [56]. However, it is also only a 2-approximation algorithm and a sample worst case instance for it (a MIN-MAX REGRET SHORTEST PATH instance) is depicted in Fig. 6.3. Note that algorithm AMU may return any of the three possible paths. But the maximum regret of the best path equals 1, while the maximum regret of the worst path equals 2. A similar example for MIN-MAX REGRET MINIMUM SPANNING TREE can be found in [56].

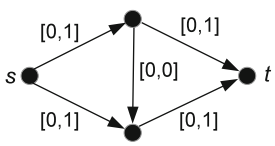


Fig. 6.3: A worst case instance for algorithm AMU

In the following we give a brief exposition of the known approaches to deal with the NP-hard minmax regret problems (see also Table 6.3). For the class

of network problems there exists a compact mixed integer programming (MIP) formulation [5, 40], which can be solved by means of some available software such as CPLEX. Another popular approach is to apply the Benders decomposition technique or a specialized branch and bound (cut) method. The detailed description of the computational tests for various instances can be found in the references given in Table 6.3.

The exact methods seem to be particularly efficient for the minmax regret version of the SHORTEST PATH problem, as they allow us to solve large problems in reasonable time (see e.g. [69]). The exact methods perform much worse for the minmax regret version of MINIMUM SPANNING TREE, which is a very interesting problem still requiring more deep investigation. The largest instances which can be solved to optimality are composed of networks having up to 40 nodes [56, 81]. For this problem a local search method seems to be more efficient. There is a very natural definition of a neighborhood of a given spanning tree. Namely, we get a neighbor X' of a spanning tree X by performing the operation $X' = X \cup \{e\} \setminus \{f\}$, where $e \in E \setminus X$ and $f \in X$. We can then apply a simple iterative improvement or more sophisticated tabu search algorithm to compute a solution. The computational tests performed in [56] suggest that the obtained solutions are close to the optimum even for large instances. Interestingly, a local minimum with respect to the specified neighborhood can also be a factor of 2 away from the global minimum, even when one starts from a solution computed by AMU (see [56]).

For the minmax regret version of MINIMUM SPANNING TREE another interesting result has been established in [33]. It turns out that the problem complexity depends on the number of intersecting intervals. Indeed, an optimal minmax regret spanning tree can be found in $O(2^k n \log n)$ time, where k is the maximum number of intervals that intersect at least one other interval. So, from this point of view, the hardest instances are the ones in which all the cost intervals are the same, for instance equal to [0,1] (a MIP approach is very poor in this case [56]). This special case is equivalent to the strongly NP-hard CENTRAL SPANNING TREE problem [6, 16]. Observe that for this problem algorithm AMU may return any solution and designing an approximation algorithm with a performance ratio better than 2 is an interesting and important open problem.

In the existing literature some other problems have been also investigated and, in the following, we briefly describe them. In [71] the minmax regret version of the TRAVELING SALESPERSON problem with interval costs and in [80] the minmax regret version of the SET COVERING problem with interval costs have been studied. Both problems are quite challenging as their deterministic versions are strongly NP-hard. The solution methods proposed in [71, 80] (a branch and cut algorithm, Benders decomposition and some heuristics) are general and can easily be extended to other minmax regret problems with interval data, whose deterministic counterparts are NP-hard. In [26] the minmax regret version of the MINIMUM SPANNING ARBORESCENCE problem with interval costs has been examined. An *arborescence* is a subgraph of a given graph G in which there is exactly one path from a given root node r to any other node of G. For undirected graphs the problem is equivalent to MINIMUM SPANNING TREE, so its minmax regret version is

strongly NP-hard. However, for acyclic directed graphs this problem can be solved in polynomial time [26].

6.3 Extensions of the Minmax Approach

In this section we introduce several extensions of the traditional minmax approach presented in the previous section. These extensions allow us to overcome some drawbacks of this approach. Namely, we will be able to model an attitude of decision makers towards risk and take additional information associated with scenario set into account.

6.3.1 Using the OWA Criterion

In decision making under uncertainty some other criteria for choosing a solution, such as minmin, Hurwicz, or Laplace (the average), are also used. For an excellent discussion on their various properties we refer the reader to [63]. It turns out that most of them are special cases of the criterion called *Weighted Ordering Averaging* aggregation (OWA for short) proposed by Yager in [85]. We will now show how to apply the OWA criterion to problem $\mathscr{P}$ under scenario set $\mathscr{U}_D$.

Let $\boldsymbol{v} = (v_1, \ldots, v_K)$ be a vector of weights, where $v_j \in [0,1]$ for each $j \in [K]$ and $v_1 + \cdots + v_K = 1$. Given a feasible solution $X \in \boldsymbol{\Phi}$, let σ be a permutation of $[K]$ such that $f(X, S_{\sigma(1)}) \geq f(X, S_{\sigma(2)}) \geq \cdots \geq f(X, S_{\sigma(K)})$. The OWA aggregation criterion is defined as follows:

$$\mathrm{OWA}(X) = \sum_{j \in [K]} v_j f(X, S_{\sigma(j)}).$$

Observe that $\mathrm{OWA}(X)$ is a convex combination of the costs $f(X,S_1), \ldots, f(X,S_K)$. Hence its value is always between the minimum and the maximum cost of X over scenarios in $\mathscr{U}_D$. In this section we assume that a vector of weights $\boldsymbol{v}$ is specified for scenario set $\mathscr{U}_D$ and we consider the following problem:

$$\textsc{Min-OWA}\ \mathscr{P}: \min_{X \in \boldsymbol{\Phi}} \mathrm{OWA}(X).$$

By fixing the weights $\boldsymbol{v}$ we get some special cases of Min-OWA $\mathscr{P}$ which are listed in Table 6.4. Observe that OWA generalizes all the basic criteria used in decision making under uncertainty except for the minmax regret (Savage) criterion.

Since Min-Max $\mathscr{P}$ is a special case of Min-Owa $\mathscr{P}$, all the negative results presented in Table 6.1 remain valid for Min-OWA $\mathscr{P}$. However, the computational properties of Min-Owa $\mathscr{P}$ strongly depend on the weight distribution in $\boldsymbol{v}$. For example, it is easily seen that Min-Min $\mathscr{P}$ and Min-Average $\mathscr{P}$ are polynomially

solvable if $\mathscr{P}$ is polynomially solvable. On the other hand, MIN-HURWICZ $\mathscr{P}$ is at least as hard as MIN-MAX $\mathscr{P}$, because it generalizes the latter problem. Table 6.5 summarizes all the known results for the MIN-OWA SHORTEST PATH problem. Most of these results remain valid for any problem $\mathscr{P}$ (see [51] for details).

Table 6.4: Special cases of MIN-OWA $\mathscr{P}$

Problem	Weights $\boldsymbol{v}$
MIN-MAX $\mathscr{P}$	$v_1 = 1$, $v_j = 0$ for $j \neq 1$
MIN-MIN $\mathscr{P}$	$v_K = 1$, $v_j = 0$ for $j \neq K$
MIN-AVERAGE $\mathscr{P}$	$v_j = \frac{1}{K}$ for all $j \in [K]$
MIN-QUANT(k) $\mathscr{P}$	$v_k = 1$, $v_j = 0$ for $j \neq k$
MIN-MEDIAN $\mathscr{P}$	$v_{\lfloor K/2 \rfloor + 1} = 1$, $v_j = 0$ for $j \neq \lfloor K/2 \rfloor + 1$
MIN-HURWICZ $\mathscr{P}$	$v_1 = \alpha$, $v_K = 1 - \alpha$, $v_j = 0$ for $j \neq 1, K$, $\alpha \in [0,1]$

Table 6.5: Summary of results for the MIN-OWA SHORTEST PATH problem

Problem	$K = 2$	$K \geq 3$ constant	K unbounded
MIN-OWA	Equivalent to MIN-HURWICZ $\mathscr{P}$	NP-hard, FPTAS	Strongly NP-hard, appr. within $v_1 K$ if the weights are non-increasing, not at all appr. if the weights are nondecreasing
MIN-MAX	NP-hard, FPTAS	NP-hard, FPTAS	Strongly NP-hard, appr. within K, not appr. within $O(\log^{1-\varepsilon} K)$, $\varepsilon > 0$
MIN-MIN	Poly. solvable	Poly. solvable	Poly. solvable
MIN-AVER.	Poly. solvable	Poly. solvable	Poly. solvable
MIN-HURWICZ	Poly. solv. if $\alpha \in [0, \frac{1}{2})$ NP-hard if $\alpha \in (\frac{1}{2}, 1]$ FPTAS if $\alpha \in (\frac{1}{2}, 1]$	NP-hard if $\alpha \in (0,1]$ FPTAS	Strongly NP-hard if $\alpha \in (0,1]$, appr. within $\alpha K + (1-\alpha)(K-2)$ if $\alpha \in [\frac{1}{2}, 1]$ $\frac{K}{\alpha}$ if $\alpha \in (0, \frac{1}{2})$, not appr. within $O(\log^{1-\varepsilon} K)$, $\varepsilon > 0$
MIN-QUANT(k)	Poly. solvable if $k = 2$ NP-hard if $k = 1$ FPTAS	Poly. solvable for $k = K$, NP-hard for constant $k \in [K-1]$, FPTAS	Strongly NP-hard for any $k \in [K-1]$, approx. within K for constant k, not at all appr. if $k = \lfloor \frac{K}{2} \rfloor + 1$

The complexity of MIN-OWA $\mathscr{P}$ depends on the weight distribution in $\boldsymbol{v}$. In particular, if the weights model the minimum or the average, then the problem is polynomially solvable when $\mathscr{P}$ is polynomially solvable. On the other hand, when $\boldsymbol{v}$ models the median, MIN-OWA SHORTEST PATH is not at all approximable. Fortunately there is a positive approximation result for the problem when the weights are non-increasing, i.e. $v_1 \geq v_2 \geq \cdots \geq v_K$. This important special case will be described in more detail in the next section.

6.3.1.1 OWA Criterion and the Robust Approach

The maximum and the average (Laplace) criteria are special cases of OWA. They form two boundary cases of non-increasing weights, i.e. when $v_1 \geq v_2 \geq \cdots \geq v_K$. We get the maximum when $v_1 = 1$, $v_j = 0$ for $j \neq 1$, and the average when $v_j = 1/K$ for all $j \in [K]$. It turns out that for non-increasing weights a general positive approximation result holds. Namely, if $\mathscr{P}$ is polynomially solvable, then MIN-OWA $\mathscr{P}$ is approximable within $v_1 K$ [51]. The idea of the approximation algorithm is to solve problem $\mathscr{P}$ for the aggregated costs $\hat{c}_i = \text{owa}_{\boldsymbol{v}}(c_i^{S_1}, \ldots, c_i^{S_K})$, $i \in [n]$. So, it generalizes the K-approximation algorithm, well known for the MIN-MAX $\mathscr{P}$ problem. Note that $v_1 \in [1/K, 1]$, so we get the worst approximation ratio when OWA is the maximum. On the other hand, when $v_1 = 1/K$, i.e. when OWA is the average, we obtain a polynomial algorithm for the problem. The assumption of non-increasing weights allows us also to construct more efficient MIP formulations for MIN-OWA $\mathscr{P}$ (see [22, 34, 76]), which makes the problem more tractable.

The non-increasing weights are compatible with the robust approach, because larger weights are assigned to larger solution costs. Furthermore, the weights allow risk-averse decision makers to model their attitude towards a risk. The more uniform is the weight distribution the less risk averse the decision maker is. In particular, extremely risk averse the decision maker will choose $v_1 = 1$, which leads to the maximum criterion and the minmax problem discussed in Sect. 6.2. Using the OWA criterion allows us to overcome another drawback of the minmax approach. When all the weights in $\boldsymbol{v}$ are positive, then the obtained solution must be Pareto optimal. We can thus reject such solutions as the one presented in Fig. 6.1a, by choosing positive (even very small) weights.

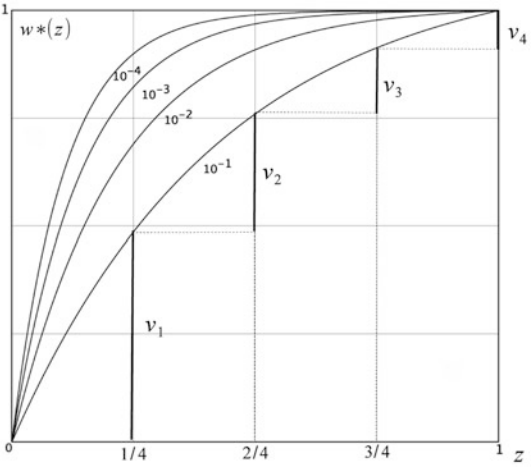
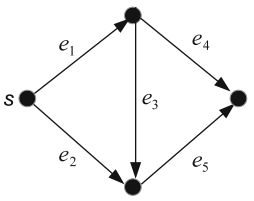


Fig. 6.4: A sample interpolation function $w^*(z) = \frac{1}{(1-\alpha)}(1-\alpha^z)$ for $\alpha \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. The weights $v_1, \ldots, v_4$ for $\alpha = 10^{-1}$

The weights $v_1, \ldots, v_K$ can be specified explicitly. However, it may be convenient to obtain them by using an *interpolation function* $w^* : [0,1] \to [0,1]$, which is assumed to be concave, nondecreasing and satisfies $w^*(0) = 0$, $w^*(1) = 1$. Having w^* we get $v_j = w^*(j/K) - w^*((j-1)/K)$ for $j \in [K]$. A sample interpolation function $w^*(z) = \frac{1}{(1-\alpha)}(1-\alpha^z)$, $\alpha \in (0,1)$, is shown in Fig. 6.4. Observe that when α tends to 0, the OWA tends to the maximum. On the other hand, when α tends to 1, the OWA tends to the average. Thus the decision maker can adjust his attitude towards a risk by fixing a single value of $\alpha \in (0,1)$. The weights obtained for $\alpha = 10^{-1}$ and $K = 4$ are also shown in Fig. 6.4. Observe that the more concave is w^* the less uniform is the weight distribution in $\boldsymbol{v}$. When w^* is a straight line, then $v_1 = \cdots = v_K$ and OWA is the average. The interpolation function $w^*(z)$ can also be defined for explicitly listed weights $v_1 \geq v_2 \geq \cdots \geq v_K$. It is enough to assume that $w^*(z)$ is the linear interpolation of the points $(0,0)$ and $(j/K, \sum_{i \leq j} v_i)$ for $j \in [K]$. Then $w^*(z)$ is a concave piecewise linear function. We use this fact in the next section, in which we describe a generalization of the OWA criterion.

6.3.2 Using the WOWA Criterion

One drawback of the OWA criterion is that it does not allow us to exploit any additional information associated with scenarios. One such important information is a probability distribution over the set $\mathcal{U}_D$. Assume that such a probability distribution is available and let p_j be the probability that scenario $S_j \in \mathcal{U}_D$ will occur, $j \in [K]$. Let us examine a sample SHORTEST PATH problem with scenario set $\mathcal{U}_D$, shown in Fig. 6.5.



p_j	0.5	0.2	0.2	0.1	
	S_1	S_2	S_3	S_4	Exp.
$X_1=\{e_1,e_4\}$	10	1	1	2	**5.6**
$X_2=\{e_1,e_3,e_5\}$	5	5	7	8	**5.7**
$X_3=\{e_2,e_5\}$	6	6	6	6	**6**

Fig. 6.5: A sample shortest path problem with four scenarios $S_1 = (5,6,0,5,0)$, $S_2 = (1,6,4,0,0)$, $S_3 = (1,6,6,0,0)$, and $S_4 = (2,6,6,0,0)$. The costs of all three paths under all scenarios are shown in the table

A natural approach to solve this problem is to choose a solution with the minimum expected cost. Hence, the path $X_1 = \{e_1, e_4\}$ is then the best choice. However, X_1 may be unreasonable for some risk averse decision makers. Observe that the probability that the path X_1 will have a large cost equal to 10 is equal to 0.5. This choice may be questionable if path X_1 is to be used only once, i.e. when a decision

is not repetitious in the same environment. On the other hand, path $X_3 = \{e_2, e_5\}$ has the smallest maximum cost and should be chosen when the minmax criterion is used and the probabilities of scenarios are ignored. Notice that the path X_3 has a deterministic cost equal to 6. However, some decision makers may feel that path $X_2 = \{e_1, e_3, e_5\}$ is better, since the probability that the cost of X_2 will be less than 6 equals 0.7 and the probability that X_2 will have a large cost, equal to 8, is only 0.1.

The sample problem demonstrates that there is a need of criterion which establishes a link between the stochastic and robust approach when scenario probabilities are available. Such a criterion can be proposed by a generalization of OWA. In order to introduce this criterion it is convenient to use the interpolation function $w^*(z)$ defined in the previous section. Consider a solution X and let σ be a permutation of $[K]$ such that $f(X, S_{\sigma(1)}) \geq \cdots \geq f(X, S_{\sigma(K)})$. The permutation σ defines also the order of scenario probabilities $p_{\sigma(1)} \geq \cdots \geq p_{\sigma(K)}$. In particular $p_{\sigma(1)}$ is the probability that the worst scenario will occur and $p_{\sigma(K)}$ is the probability that the best scenario will occur for X. Define now the weights $\omega_j = w^*(\sum_{i \leq j} p_{\sigma(i)}) - w^*(\sum_{i<j} p_{\sigma(i)})$, $j \in [K]$ and let

$$\text{WOWA}(X) = \sum_{j \in [K]} \omega_j f(X, S_{\sigma(j)}).$$

We have thus obtained the Weighted OWA criterion (WOWA for short), first proposed in [83]. A trivial verification shows that $\omega_j \in [0,1]$ for all $j \in [K]$ and $\omega_1 + \cdots + \omega_K = 1$. The value of ω_j can be seen as a *distorted probability* of scenario $S_{\sigma(j)}$. The value of ω_j depends on p_j and the rank position of scenario $S_{\sigma(j)}$ for solution X, so it is solution dependent. Hence, WOWA(X) can be seen as the expected cost of solution X with respect to the distorted probabilities. For a more detailed interpretation of this expectation we refer the reader to [29].

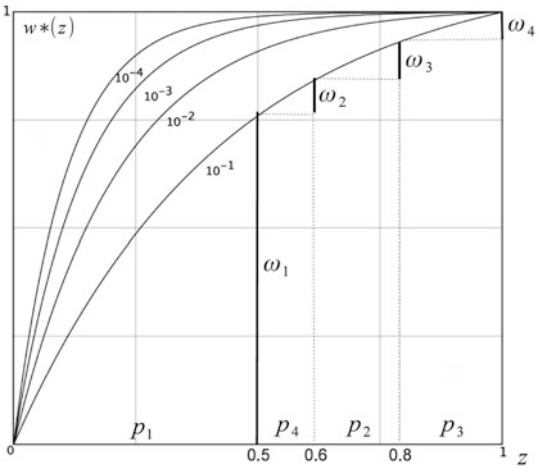


Fig. 6.6: Computing the weights $\omega_1, \ldots, \omega_4$ for path $\{e_1, e_4\}$ in Fig. 6.4

Let us look at the sample problem in Fig. 6.5 again. For path $X_1 = \{e_1, e_4\}$ we have $\sigma = (1,4,2,3)$. The computations of the weights $\omega_1, \ldots, \omega_4$ is shown in Fig. 6.6. One can see in Fig. 6.6 how scenario probabilities are distorted. For example, $\omega_1 > p_1$ and $\omega_4 < p_3$, so we assign larger probability to the worst scenario S_1 and smaller to good scenario S_3. If the probability distribution in $\mathscr{U}_D$ is uniform, then WOWA becomes OWA, because $\omega_j = v_j$ for each $j \in [K]$. The uniform probability distribution results from applying the *principle of insufficient reason*, i.e. in a situation under uncertainty, when it is not possible to distinguish more or less probable scenarios [63]. If $w^*(z)$ is a straight line, or equivalently, $v_1 = v_2 = \cdots = v_K$, then WOWA becomes the expected value, because $\omega_j = p_{\sigma(j)}$ is then just the scenario probability. We thus can see that WOWA is a very general criterion. It contains both OWA and the expected value as special cases. It allows us to establish a link between the robust and stochastic approaches.

Since MIN-MAX $\mathscr{P}$ is a special case of MIN-WOWA $\mathscr{P}$, all the negative results shown in Table 6.1 remain true for the latter problem. Fortunately, when $w^*(z)$ is the linear interpolation function for the nonincreasing weights $v_1 \geq v_2 \geq \cdots \geq v_K$ (see Sect. 6.3.1), then the problem is approximable within $v_1 K$ if $\mathscr{P}$ is polynomially solvable [54]. An idea is to solve $\mathscr{P}$ for the aggregated cots $\hat{c}_i = \text{wowa}_{\boldsymbol{v},\boldsymbol{p}}(c_i^{S_1}, \ldots, c_i^{S_K})$, $i \in [n]$. Note that this approximation ratio is the same as for MIN-OWA $\mathscr{P}$. For the linear interpolation function $w^*(z)$ a compact MIP formulation for the problem can also be constructed [54, 77]. Some computational tests for the MIP formulation and the approximation algorithm were performed in [54].

6.4 Robust Optimization with Incremental Recourse

In this section we address the adjustable approach to combinatorial optimization problems with uncertain element costs, introduced in [72], called the *robust optimization with incremental recourse*. It extends the concept of robustness to deal with uncertainties by incorporating adjustable actions, after an element cost scenario is realized. Namely, the decision maker chooses first the best initial solution, taking into account that a worst scenario can happen (the first stage). Then he makes some incremental changes in the initial solution chosen, subject to a given distance measure, in order to obtain another one (the incremental recourse stage).

Formally, the robust optimization problem $\mathscr{P}$ with incremental recourse can be stated as follows:

$$\text{ROIR } \mathscr{P}: \min_{X \in \boldsymbol{\Phi}} (f(X) + \max_{S \in \mathscr{U}} \min_{Y \in \boldsymbol{\Phi}_X^k} f(Y,S)),$$

where $f(X) = \sum_{e_i \in X} C_i$ is the cost of an initial solution X and $\boldsymbol{\Phi}_X^k = \{Y \in \boldsymbol{\Phi} : d(X,Y) \leq k\}$ is the *incremental set*, i.e. the set of possible solutions in the incremental

recourse stage, where $d(X,Y)$ is a fixed measure of the distance between the initial solution X and the incremental solution Y. The distance $d(X,Y)$ is also called an *incremental function* bounded by a specified parameter k. Finally, $f(Y,S) = \sum_{e_i \in Y} c_i^S$ is the cost of solution Y under scenario S. The most popular distance measures $d(X,Y)$, proposed in literature [17, 18, 72, 82], are: the *element inclusion distance* $d(X,Y) = |Y \setminus X|$, the *element exclusion distance* $d(X,Y) = |X \setminus Y|$, and the *element symmetric difference distance* $d(X,Y) = |X \oplus Y|$. It is worthwhile to mention that for MINIMUM SPANNING TREE the above distance measures are equivalent from the computational point of view.

The concept of the robustness with incremental recourse is similar in spirit to the one of the *recoverable robustness*, proposed in [62] for linear programing under uncertainty. In the recoverable approach limited recovery actions are permitted after uncertain parameters reveal. Later, in [17–19] the recoverable robustness, called k-distance recoverable robustness, has been applied to some classical combinatorial optimization problems. Yet another interesting concept of recoverable robustness, proposed in [17, 18], is the *rent recoverable robustness* in which, in the second recoverable stage, the number of elements that can be replaced is not limited, but deviating from previous choice comes at extra cost.

The ROIR $\mathscr{P}$ problem contains the following three inner problems. The first one is the *incremental problem*:

$$\text{INC}(X,S^*)\ \mathscr{P}:\ \min_{Y \in \Phi_X^k} f(Y,S^*),$$

where we are given an initial solution X and a cost scenario S^* revealed. We wish to make incremental changes in X, subject to the constraint $d(X,Y) \leq k$, which lead to the maximum improvement in the objective function. This problem is a special case of ROIR $\mathscr{P}$. Indeed, it is sufficient to set $\mathscr{U} = \{S^*\}$, the initial costs $C_i = 0$ if $e_i \in X$ and M otherwise, where M is a sufficiently large number, for example $M \geq nC$, $C = max_{e_i \in E}\{c_i^{S^*}\}$. Several incremental versions of network problems have been investigated in [82]. The second inner problem is the *adversarial* one:

$$\text{ADV}(X)\ \mathscr{P}:\ \max_{S \in \mathscr{U}} \min_{Y \in \Phi_X^k} f(Y,S).$$

In this problem we seek a scenario $S \in \mathscr{U}$ that maximizes $\text{INC}(X,S)$ with respect to a given solution X. The problem ROIR $\mathscr{P}$ reduces to $\text{ADV}(X)$ $\mathscr{P}$ when we set $C_i = 0$ if $e_i \in X$ and $C_i = M$, otherwise. The last inner problem is MIN-MAX $\mathscr{P}$. We get this problem after fixing $C_i = 0$ for each $e_i \in E$ and $k = 0$, which implies $\Phi_X^k = \{X\}$. We thus can see that ROIR $\mathscr{P}$ generalizes MIN-MAX $\mathscr{P}$. In consequence, all the negative results for MIN-MAX $\mathscr{P}$ remain valid for the robust incremental version of $\mathscr{P}$.

In the next two sections we will present the known results on ROIR $\mathscr{P}$. We will show that the complexity of this problem highly depends on both the uncertainty representation and the distance measure. As we will see, there are a lot of things to do in this area. In particular, there is lack of approximation algorithms for the considered problem. The most of presented results are negative ones.

6.4.1 Discrete Uncertainty Representation

We now give a brief summary of the complexity results on ROIR $\mathscr{P}$ under the discrete scenario uncertainty, i.e. when $\mathscr{U} = \mathscr{U}_D$. In this case, the only complexity results that exist in the literature are for the element inclusion distance $d(X,Y) = |Y \setminus X|$. Unfortunately, all of them are negative ones, except for the robust incremental problems with one scenario (see Table 6.6).

Table 6.6: Complexity results for ROIR $\mathscr{P}$ with scenario set $\mathscr{U}_D$ and the element inclusion distance

ROIR $\mathscr{P}$	Constant K	Unbounded K
SHORTEST PATH	Strongly NP-hard, not at all appr. for $K = 1$ and $k \geq 2$ [18]	
MINIMUM SPANNING TREE	NP-hard in sp-graphs for $K = 2$ and constant k [58]	Strongly NP-hard, not at all appr. for unbounded k [58]
SELECTION	NP-hard for $K = 2$ and $k \geq 1$, solv. in $O((p-k+1)n^2)$ for $K = 1$ [53]	Strongly NP-hard, not at all appr. for any const. $k \geq 1$ [53]
MINIMUM MATROID BASE	Poly. solvable for constant k and for $K = 1$ [17]	

Since MIN-MAX $\mathscr{P}$ is a special case of ROIR $\mathscr{P}$, all negative results presented in Table 6.1 for MIN-MAX $\mathscr{P}$ are still true for ROIR $\mathscr{P}$. Hence, Table 6.6 can be completed by complexity results for other combinatorial problem. One can observe (see Table 6.6) that ROIR $\mathscr{P}$ can be much harder than its minmax counterpart. For instance, the ROIR SHORTEST PATH problem is strongly NP-hard and not at all approximable even for one scenario. Notice that MIN-MAX SHORTEST PATH in this case is a deterministic problem and so it is polynomially solvable. Let us also mention a more general result on KNAPSACK that has been examined in [19], i. e. the problem under the discrete uncertainty in the objective and the constraint with a distance measure that takes into account the element inclusion and exclusion distances. The problem turned out to be inapproximable for unbounded K, but pseudopolynomially solvable for constant K.

We now look into the adversarial problem with scenario set $\mathscr{U}_D$. It easily seen that the complexity of this problem highly relies on the complexity of the incremental problem. Indeed, solving ADV(X) $\mathscr{P}$, for a given initial solution X, boils down to solving INC(X,S) $\mathscr{P}$ for every $S \in \mathscr{U}_D$ and choosing a scenario which results in the maximum cost. It turns out (see [82]) that the incremental versions of SHORTEST PATH and MINIMUM SPANNING TREE, with the element inclusion distance function, are polynomially solvable, and the incremental version of MINIMUM ASSIGNMENT can be solved in random polynomial time. Unfortunately the incremental versions of MINIMUM S-T CUT and SHORTEST PATH, with the element symmetric difference distance function, $d(X,Y) = |X \oplus Y|$, and the incremental version of SHORTEST PATH, with the element exclusion distance, $d(X,Y) = |X \setminus Y|$, are

NP-hard [72, 82]. In consequence, their adversarial and robust incremental versions are also NP-hard.

6.4.2 Interval Uncertainty Representation

In this section we are concerned with ROIR $\mathscr{P}$ under the interval uncertainty representation and with three distance measures. We start by showing a few complexity results for the case $\mathscr{U} = \mathscr{U}_I$ and the element inclusion distance $d(X,Y) = |Y \setminus X|$ (see Table 6.7). To the authors' knowledge, nothing more has been recorded in the literature on the robust incremental optimization with recourse under the scenario set $\mathscr{U}_I$.

Table 6.7: Complexity results for ROIR $\mathscr{P}$ problems with scenario set $\mathscr{U}_I$ and the element inclusion distance

ROIR $\mathscr{P}$	Complexity
SHORTEST PATH	Strongly NP-hard, not at all appr., poly. solvable in sp-graphs [18]
SELECTION	Solvable in $O((p-k+1)n^2)$ time [53]
MINIMUM MATROID BASE	Poly. solvable for constant k [17]

In the robust incremental optimization with recourse there is a link between the interval uncertainty representation $\mathscr{U}_I$ and the discrete one $\mathscr{U}_D$, namely, the ROIR $\mathscr{P}$ problem with scenario set $\mathscr{U}_I$ can be rewritten as follows:

$$\begin{aligned}\min_{X\in\Phi}\Big(\sum_{e_i\in X} C_i + \max_{S\in\mathscr{U}_I}\min_{Y\in\Phi_X^k}\sum_{e_i\in Y} c_i^S\Big) &= \min_{X\in\Phi}\Big(\sum_{e_i\in X} C_i + \min_{Y\in\Phi_X^k}\sum_{e_i\in Y}(c_i+d_i)\Big)\\ &= \min_{X\in\Phi}\Big(\sum_{e_i\in X} C_i + \mathrm{INC}\big(X,(c_i+d_i)_{i\in[n]}\big)\Big).\end{aligned}$$

From the above it follows that ROIR $\mathscr{P}$ with scenario set $\mathscr{U}_I$ is equivalent to ROIR $\mathscr{P}$ with only one scenario $S = (c_i+d_i)_{i\in[n]}$. This property has been exploited in construction of the algorithms for the SHORTEST PATH, SELECTION and MINIMUM MATROID BASE problems with the element inclusion distance [17, 18, 53]. Moreover, using it we can conclude that ROIR KNAPSACK is at least NP-hard. One can also deduce from the NP-hardness of the incremental versions of SHORTEST PATH and MINIMUM S-T CUT with the element symmetric difference and exclusion distances the NP-hardness of their robust incremental counterparts [72, 82] (see Sect. 6.4.1).

Let us now discuss the interval uncertainty representation which allows us to control the amount of uncertainty. Following [72], we define two scenario sets. Namely, given $\ell > 0$

$$\mathscr{U}_{I1}^{\ell} = \{S = (c_i^S)_{i\in[n]} : c_i^S = c_i + \delta_i, 0 \le \delta_i \le d_i, \sum_{i\in[n]} \delta_i \le \ell\},$$
$$\mathscr{U}_{I2}^{\ell} = \{S = (c_i^S)_{i\in[n]} : c_i^S = c_i + \delta_i d_i, \delta_i \in \{0,1\}, \sum_{i\in[n]} \delta_i \le \ell\}.$$

It easy to see that $\mathscr{U}_{I1}^{\ell}$ models the situation where the total amount of deviation in the element costs is bounded by a specified ℓ. The set $\mathscr{U}_{I2}^{\ell}$ is the set of extreme points (scenarios) of $\mathscr{U}_I^{\ell}$, here $\ell \in [0,n]$ (for $\ell = n$ it is the set of extreme points of $\mathscr{U}_I$).

We first review of the complexity results on the robust incremental optimization problems with recourse under the scenario set $\mathscr{U}_{I1}^{\ell}$. In this case the adversarial version of SHORTEST PATH with the element inclusion distance can be formulated as a linear program and, in consequence, is solvable in polynomial time [72]. Unfortunately, the robust incremental version of SHORTEST PATH with the same distance is strongly NP-hard and not approximable within a factor of 2 even for $\ell = k = 1$. The rest of the hardness results for the adversarial and robust incremental versions of SHORTEST PATH with the element symmetric difference and exclusion distances presented in [72] follow from the NP-hardness of its incremental versions [82] (see Sect. 6.4.1). The adversarial version of MINIMUM SPANNING TREE with the element inclusion distance is polynomially solvable since it can be also modeled as a linear program [72]. However, the complexity status of the robust incremental version of MINIMUM SPANNING TREE remains open.

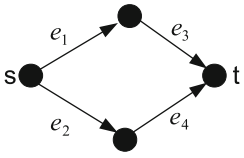
The complexity situation under the scenario set $\mathscr{U}_{I2}^{\ell}$ is much worse that under the set $\mathscr{U}_{I1}^{\ell}$. The adversarial version of SHORTEST PATH with all three distance measures considered is NP-hard and not approximable within a factor of 2 [72]. The robust incremental version of SHORTEST PATH with the element inclusion distance is strongly NP-hard and not approximable within a factor of 2 even for $\ell = k = 1$ [72]. This result improves the one shown in [18], where it has been proved that the problem is NP-hard for constant $k \ge 1$. For the remaining two distance measures, the robust incremental version of SHORTEST PATH is NP-hard and not approximable within a factor of 2, due to the hardness of the adversarial counterpart [72]. The adversarial version of MINIMUM SPANNING TREE with the element inclusion distance is NP-hard and hence its robust incremental version is NP-hard as well [72].

6.5 Robust Two-Stage Problems

In many applications, the discrete optimization problem has a two-stage nature. Namely, a partial solution is formed in the first stage, when the element costs are precisely known. This partial solution is then completed optimally after a true scenario reveals. Let C_i be the deterministic, first stage cost of element $e_i \in E$ and let c_i^S be the second stage cost of element e_i under scenario $S \in \mathscr{U}$. In this section we study the following problem:

$$\text{TWO-STAGE } \mathscr{P} : \min_{X \subseteq E}\Big(\sum_{e_i \in X} C_i + \max_{S \in \mathscr{U}} \min_{\{Y \subseteq E:\, X \cup Y \in \Phi\}} \sum_{e_i \in Y} c_i^S\Big).$$

Note that a solution to this problem is determined by a subset X of the elements, chosen in the first stage. Given X and scenario S we compute Y such that $X \cup Y \in \Phi$. It may happen that X cannot be completed to any solution from Φ. In this case we assume that the cost of X is infinite. Given Φ, let us define set Φ' in the following way: $X' \in \Phi'$ if there is $X \in \Phi$ such that $X \subseteq X'$. Hence $\Phi \subseteq \Phi'$ and Φ' contains all solutions from Φ and all the supersets of these solutions. In the one-stage robust problems described in Sect. 6.2, an optimal solution is the same when we replace Φ with Φ'. However, for the examined two-stage model the problems with Φ and Φ' may be quite different. Consider the sample TWO-STAGE SHORTEST PATH problem with two scenarios, shown in Fig. 6.7. If Φ contains two paths $\{e_1, e_3\}$ and $\{e_2, e_4\}$, then we can choose either e_1 or e_2 in the first stage. In both cases the maximum cost of the obtained path after the second stage equals $M + 1$. However, if we replace Φ with Φ', then we can choose both e_1 and e_2 in the first stage and the maximum cost of the obtained solution after the second stage equals only 3. This example demonstrates that it may be profitable to add some redundant elements in the first stage. It is thus justified to explore the complexity of the problem with both Φ and Φ'.



	C_i	S_1	S_2
e_1	1	M	M
e_2	1	M	M
e_3	M	M	1
e_4	M	1	M

Fig. 6.7: A sample TWO-STAGE SHORTEST PATH problem with two scenarios

6.5.1 Discrete Uncertainty Representation

The known complexity results for the TWO-STAGE $\mathscr{P}$ problem, when $\mathscr{U} = \mathscr{U}_D$ are shown in Table 6.8. All the basic problems are NP-hard even for two scenarios. One exception is the two-stage version of MINIMUM SPANNING TREE, for which no negative result for constant K has appeared in the literature yet. As usual, the problems become more complex when the number of scenarios is a part of input (see Table 6.8). The negative approximation results for the two-stage versions of SHORTEST PATH, MINIMUM S-T CUT and MINIMUM ASSIGNMENT have been established in [55] by showing a cost preserving reduction from the MIN-MAX REPRESENTATIVES SELECTION problem. For the two-stage versions of MINIMUM SPANNING

TREE, SELECTION, and MINIMUM ASSIGNMENT some approximation algorithms have been recently proposed [48, 53, 60]. They are based on randomized rounding of LP programs, which is a promising technique to construct approximation algorithms for robust problems with discrete scenario sets.

Table 6.8: Complexity results for various TWO-STAGE $\mathscr{P}$ problems with scenario set $\mathscr{U}_D$

TWO-STAGE $\mathscr{P}$	Constant K	Unbounded K
SHORTEST PATH	NP-hard for $K=2$ [55]	Strongly NP-hard [55], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [55]*,
MINIMUM SPANNING TREE		Strongly NP-hard [48], not appr. within $(1-\varepsilon)\log n$ for any $\varepsilon > 0$ [48], appr. within $O(\log^2 n)$ with high probability [48]
MINIMUM S-T CUT	NP-hard for $K=2$ [55]	Strongly NP-hard [55], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [55]*,
MINIMUM ASSIGNMENT	NP-hard for $K=2$ [55]	Strongly NP-hard [55], not appr. within $O(\log^{1-\varepsilon} K)$ for any $\varepsilon > 0$ [55]*, appr. within $1/\beta$, $\beta \in (0,1)$ to match at least $n(1-\beta)$ nodes [60]
SELECTION	NP-hard for $K=2$ [13]	Strongly NP-hard [53], not appr. within $(1-\varepsilon)\log n$ for any $\varepsilon > 0$ [53], appr. within $O(\log K + \log n)$ with high probability [53]

The symbol * means that the negative result holds only for Φ (we do not know if it holds for Φ')

6.5.2 Interval Uncertainty Representation

Let us first deal with scenario set $\mathscr{U}_I = \Pi_{i\in[n]}[c_i, c_i + d_i]$. In this case, the two-stage problem can be rewritten as follows:

$$\min_{\{X,Y\subseteq E:\, X\cup Y\in\Phi\}} \Big(\sum_{e_i\in X} C_i + \sum_{e_i\in Y}(c_i+d_i)\Big). \tag{6.4}$$

It follows easily that an optimal solution to (6.4) can be obtained by solving the deterministic problem $\mathscr{P}$ for the element costs $\hat{c}_i = \min\{C_i, c_i + d_i\}$, $i \in [n]$. If Z is an optimal solution to this problem, then for each $e_i \in Z$ we choose e_i in the first stage when $C_i \leq c_i + d_i$ and we choose e_i in the second stage otherwise.

Consider now the more general scenario set $\mathscr{U}_I^\ell$. If all the first stage costs C_i, $i \in [n]$, are large enough, then $X = \emptyset$ and the two-stage problem reduces to a special case of the adversarial problem, where $\Phi_X^k = \Phi$, (see Sect. 6.4.1), namely:

$$\max_{S \in \mathscr{U}_I^\ell} \min_{Y \in \Phi} \sum_{e_i \in Y} c_i^S. \tag{6.5}$$

It turns out (see [55]) that (6.5) is strongly NP-hard when $\mathscr{P}$ is MINIMUM SPANNING TREE or SHORTEST PATH. This fact immediately implies that the two-stage versions of both problems, under scenario set $\mathscr{U}_I^\ell$, are strongly NP-hard. It is worth pointing out that the corresponding MIN-MAX $\mathscr{P}$ problem, which can be obtained by interchanging the min and max operators in (6.5), is polynomially solvable (see Sect. 6.2.1.2).

There is a number of interesting open questions related to TWO-STAGE $\mathscr{P}$ with scenario set $\mathscr{U}_I^\ell$. We do not know if the problem is NP-hard when ℓ is constant (only the boundary cases $\ell = 1$ and $\ell = n$ are known to be polynomially solvable). There is also lack of positive results for this problem, in particular approximation algorithms with some guaranteed worst case ratio.

6.6 Conclusions

In this chapter we have described a class of robust discrete optimization problems with uncertain costs. We have discussed two most popular methods of modeling the uncertainty, namely the discrete and interval uncertainty representations. A lot of results and new concepts in this area have appeared in literature since 1997, when the book [61] has been published. In particular, the complexity of basic minmax (regret) problems, described in [61], has been explored more deeply. Unfortunately, with a few exceptions, all these problems are NP-hard and solving them is often a challenging task. There is still a number of important open problems in this area. One of them is to decide whether the approximation ratio of 2 is the best possible for the minmax regret problems with interval data. There is also lack of positive results for the minmax regret problems under the discrete uncertainty representation. The minmax approach has been recently generalized by using the OWA and WOWA criteria, which allow us to take both the attitude of decision makers towards a risk and scenario probabilities into account. Some special cases, for instance the problems with the Hurwicz criterion, still require more deep investigation.

In this chapter we have also reviewed some recent extensions of the minmax approach, in which computing an optimal solution is a two-stage process. In the robust incremental recourse approach an initial solution can be modified to some extent after observing a true scenario. In the two-stage approach a solution is built in two stages. A part of this solution is constructed in the first stage and the rest is constructed after a true scenario reveals. Most results, known on both approaches, are negative and there are many open questions related to their complexity and approximability.

Acknowledgements This work was partially supported by the National Center for Science (Narodowe Centrum Nauki), grant 2013/09/B/ST6/01525.

References

1. Ahuja, R.K., Magnanti, T.L., Orlin, J.B.: Network Flows: Theory, Algorithms, and Applications. Prentice Hall, Englewood Cliffs (1993)
2. Aissi, H., Bazgan, C., Vanderpooten, D.: Complexity of the min-max and min-max regret assignment problems. Oper. Res. Lett. **33**, 634–640 (2005)
3. Aissi, H., Bazgan, C., Vanderpooten, D.: Approximation of min-max and min-max regret versions of some combinatorial optimization problems. Eur. J. Oper. Res. **179**, 281–290 (2007)
4. Aissi, H., Bazgan, C., Vanderpooten, D.: Complexity of the min-max (regret) versions of min cut problems. Discret. Optim. **5**, 66–73 (2008)
5. Aissi, H., Bazgan, C., Vanderpooten, D.: Min-max and min-max regret versions of combinatorial optimization problems: a survey. Eur. J. Oper. Res. **197**, 427–438 (2009)
6. Aron, I.D., van Hentenryck, P.: On the complexity of the robust spanning tree problem with interval data. Oper. Res. Lett. **32**, 36–40 (2004)
7. Aron, I.D., van Hentenryck, P.: A constraint satisfaction approach to the robust spanning tree problem with interval data. CoRR (2013), abs/1301.0552
8. Averbakh, I.: Minmax regret solutions for minimax optimization problems with uncertainty. Oper. Res. Lett. **27**, 57–65 (2000)
9. Averbakh, I.: On the complexity of a class of combinatorial optimization problems with uncertainty. Math. Program. **90**, 263–272 (2001)
10. Averbakh, I.: Computing and minimizing the relative regret in combinatorial optimization with interval data. Discret. Optim. **2**, 273–287 (2005)
11. Averbakh, I., Lebedev, V.: Interval data minmax regret network optimization problems. Discret. Appl. Math. **138**, 289–301 (2004)
12. Averbakh, I., Lebedev, V.: On the complexity of minmax regret linear programming. Eur. J. Oper. Res. **160**, 227–231 (2005)
13. Baumann, F., Buchheim, C., Ilyina, A.: A Lagrangean decomposition approach for robust combinatorial optimization. Technical Report, Optimization Online (2014)
14. Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: Robust Optimization. Princeton Series in Applied Mathematics. Princeton University Press, Princeton (2009)
15. Bertsimas, D., Sim, M.: Robust discrete optimization and network flows. Math. Program. **98**, 49–71 (2003)
16. Bezrukov, S.L., Kaderali, F., Poguntke, W.: On central spanning trees of a graph. In: Combinatorics and Computer Science, 8th Franco-Japanese and 4th Franco-Chinese Conference. Lecture Notes in Computer Science, vol. 1120, pp. 53–57. Springer, Berlin (1995)
17. Büsing, C.: Recoverable robustness in combinatorial optimization. Ph.D. thesis, Technical University of Berlin, Berlin (2011)
18. Büsing, C.: Recoverable robust shortest path problems. Networks **59**, 181–189 (2012)
19. Büsing, C., Koster, A.M.C.A., Kutschka, M.: Recoverable robust knapsacks: the discrete scenario case. Optim. Lett. **5**, 379–392 (2011)
20. Catanzaro, D., Labbé, M., Salazar-Neumann, M.: Reduction approaches for robust shortest path problems. Comput. Oper. Res. **38**, 1610–1619 (2011)
21. Chanas, S., Zieliński, P.: The computational complexity of the criticality problems in a network with interval activity times. Eur. J. Oper. Res. **136**, 541–550 (2002)
22. Chassein, A.B., Goerigk, M.: Alternative formulations for the ordered weighted averaging objective. Inf. Process. Lett. **115**, 604–608 (2015)
23. Chassein, A.B., Goerigk, M.: A new bound for the midpoint solution in minmax regret optimization with an application to the robust shortest path problem. Eur. J. Oper. Res. **244**, 739–747 (2015)

24. Conde, E.: An improved algorithm for selecting p items with uncertain returns according to the minmax regret criterion. Math. Program. **100**, 345–353 (2004)
25. Conde, E.: A 2-approximation for minmax regret problems via a mid-point scenario optimal solution. Optim. Lett. **38**, 326–327 (2010)
26. Conde, E., Candia, A.: Minimax regret spanning arborescences under uncertain costs. Discret. Optim. **182**, 561–577 (2007)
27. Deineko, V.G., Woeginger, G.J.: Pinpointing the complexity of the interval min-max regret knapsack problem. Discret. Optim. **7**, 191–196 (2010)
28. Deineko, V.G., Woeginger, G.J.: Complexity and in-approximability of a selection problem in robust optimization. 4OR: Q. J. Oper. Res. **11**, 249–252 (2013)
29. Diecidue, E., Wakker, P.P.: On the intuition of rank-dependent utility. J. Risk Uncertain. **23**, 281–298 (2001)
30. Doerr, B.: Improved approximation algorithms for the min-max selecting items problem. Inf. Process. Lett. **113**, 747–749 (2013)
31. Dolgui, A., Kovalev, S.: Min-max and min-max (relative) regret approaches to representatives selection problem. 4OR: Q. J. Oper. Res. **10**, 181–192 (2012)
32. Dubois, D., Fortemps, P.: Computing improved optimal solutions to max-min flexible constraint computing improved optimal solutions to max-min flexible constraint satisfaction problems. Eur. J. Oper. Res. **118**, 95–126 (1999)
33. Escoffier, B., Monnot, J., Spanjaard, O.: Some tractable instances of interval data minmax regret problems. Oper. Res. Lett. **36**, 424–429 (2008)
34. Fernandez, E., Pozo, M.A., Puerto, J.: Ordered weighted average combinatorial optimization: formulations and their properties. Discret. Appl. Math. **169**, 97–118 (2014)
35. Fortin, J., Zieliński, P., Dubois, D., Fargier, H.: Criticality analysis of activity networks under interval uncertainty. J. Sched. **13**, 609–627 (2010)
36. Furini, F., Iori, M., Martello, S., Yagiura, M.: Heuristic and exact algorithms for the interval min-max regret knapsack problem. Informs J. Comput. **27**, 392–405 (2015)
37. Garey, M.R., Johnson, D.S.: Computers and Intractability. A Guide to the Theory of NP-Completeness. W.H. Freeman and Company, New York (1979)
38. Goerigk, M., Shőbel, A.: Algorithm engineering in robust optimization. In: Kliemann, L., Sanders, P. (eds.) Algorithm Engineering: Selected Results and Surveys. Lecture Notes in Computer Science. State of the Art. Springer **9220** (2016, to appear)
39. Karasan, O., Pinar, M., Yaman, H.: The robust shortest path problem with interval data. Technical Report, Bilkent University, Ankara (2001)
40. Kasperski, A.: Discrete Optimization with Interval Data - Minmax Regret and Fuzzy Approach. Studies in Fuzziness and Soft Computing, vol. 228. Springer, Berlin (2008)
41. Kasperski, A., Zieliński, P.: An approximation algorithm for interval data minmax regret combinatorial optimization problems. Inf. Process. Lett. **97**, 177–180 (2006)
42. Kasperski, A., Zieliński, P.: The robust shortest path problem in series-parallel multidigraphs with interval data. Oper. Res. Lett. **34**, 69–76 (2006)
43. Kasperski, A., Zieliński, P.: On combinatorial optimization problems on matroids with uncertain weights. Eur. J. Oper. Res. **177**, 851–864 (2007)
44. Kasperski, A., Zieliński, P.: On the existence of an FPTAS for minmax regret combinatorial optimization problems with interval data. Oper. Res. Lett. **35**, 525–532 (2007)
45. Kasperski, A., Zieliński, P.: On the approximability of minmax (regret) network optimization problems. Inf. Process. Lett. **109**, 262–266 (2009)
46. Kasperski, A., Zieliński, P.: A randomized algorithm for the min-max selecting items problem with uncertain weights. Ann. Oper. Res. **172**, 221–230 (2009)
47. Kasperski, A., Zieliński, P.: Minmax regret approach and optimality evaluation in combinatorial optimization problems with interval and fuzzy weights. Eur. J. Oper. Res. **200**, 680–687 (2010)
48. Kasperski, A., Zieliński, P.: On the approximability of robust spanning problems. Theor. Comput. Sci. **412**, 365–374 (2011)
49. Kasperski, A., Zieliński, P.: Bottleneck combinatorial optimization problems with uncertain costs and the OWA criterion. Oper. Res. Lett. **41**, 639–643 (2013)

50. Kasperski, A., Zieliński, P.: Minmax (regret) sequencing problems. In: Werner, F., Sotskov, Y. (eds.) Sequencing and Scheduling with Inaccurate Data, chap. 8, pp. 159–210. Nova Science Publications, Inc. New York (2014)
51. Kasperski, A., Zieliński, P.: Combinatorial optimization problems with uncertain costs and the owa criterion. Theor. Comput. Sci. **565**, 102–112 (2015)
52. Kasperski, A., Zieliński, P.: Robust independent set problems on interval graphs. Optim. Lett. **9**, 427–436 (2015)
53. Kasperski, A., Zieliński, P.: Robust recoverable and two-stage selection problems (2015). http://arxiv.org/abs/1505.06893
54. Kasperski, A., Zieliński, P.: Using the WOWA operator in robust discrete optimization problems. Int. J. Approx. Reason. **68**, 54–67 (2016)
55. Kasperski, A., Zieliński, P.: Robust two-stage network problems. In: Proceedings of the OR 2015 Conference (accepted paper) (2016)
56. Kasperski, A., Makuchowski, M., Zieliński, P.: A tabu search algorithm for the minmax regret minimum spanning tree problem with interval data. J. Heuristics **18**, 593–625 (2012)
57. Kasperski, A., Kurpisz, A., Zieliński, P.: Approximating the min-max (regret) selecting items problem. Inf. Process. Lett. **113**, 23–29 (2013)
58. Kasperski, A., Kurpisz, A., Zieliński, P.: Recoverable robust combinatorial optimization problems. In: Operations Research Proceedings 2012, pp. 147–153 (2014)
59. Kasperski, A., Kurpisz, A., Zieliński, P.: Approximability of the robust representatives selection problem. Oper. Res. Lett. **43**, 16–19 (2015)
60. Katriel, I., Kenyon-Mathieu, C., Upfal, E.: Commitment under uncertainty: two-stage matching problems. Theor. Comput. Sci. **408**, 213–223 (2008)
61. Kouvelis, P., Yu, G.: Robust Discrete Optimization and its Applications. Kluwer Academic Publishers, Dordrecht (1997)
62. Liebchen, C., Lübbecke, M.E., Möhring, R.H., Stiller, S.: The concept of recoverable robustness, linear programming recovery, and railway applications. In: Robust and Online Large-Scale Optimization. Lecture Notes in Computer Science, vol. 5868, pp. 1–27. Springer, Berlin (2009)
63. Luce, R.D., Raiffa, H.: Games and Decisions: Introduction and Critical Survey. Dover, New York (1989)
64. Mastin, A., Jaillet, P., Chin, S.: Randomized minmax regret for combinatorial optimization under uncertainty. In: 26th International Symposium on Algorithms and Computation ISAAC, pp. 491–501 (2015)
65. Monaci, M., Pferschy, U.: On the robust knapsack problem. SIAM J. Optim. **23**, 1956–1982 (2013)
66. Montemanni, R.: A benders decomposition approach for the robust spanning tree problem with interval data. Eur. J. Oper. Res. **174**, 1479–1490 (2006)
67. Montemanni, R., Gambardella, L.M.: An exact algorithm for the robust shortest path problem with interval data. Comput. Oper. Res. **31**, 1667–1680 (2004)
68. Montemanni, R., Gambardella, L.M.: A branch and bound algorithm for the robust spanning tree problem with interval data. Eur. J. Oper. Res. **161**, 771–779 (2005)
69. Montemanni, R., Gambardella, L.M.: The robust shortest path problem with interval data via benders decomposition. 4OR: Q. J. Oper. Res. **3**(4), 315–328 (2005)
70. Montemanni, R., Gambardella, L.M., Donati, A.V.: A branch and bound algorithm for the robust shortest path problem with interval data. Oper. Res. Lett. **32**, 225–232 (2004)
71. Montemanni, R., Barta, J., Mastrolilli, M., Gambardella, L.M.: The robust traveling salesman problem with interval data. Transp. Sci. **41**(3), 366–381 (2007)
72. Nasrabadi, E., Orlin, J.B.: Robust optimization with incremental recourse (2013). http://arxiv.org/abs/1312.4075
73. Nikulin, Y.: Simulated annealing algorithm for the robust spanning tree problem. J. Heuristics **14**, 391–402 (2008)
74. Nobibon, F.T., Leus, R.: Complexity results and exact algorithms for robust knapsack problems. J. Optim. Theory Appl. **161**, 533–552 (2014)

75. Nobibon, F.T., Leus, R.: Robust maximum weighted independent-set problems on interval graphs. Optim. Lett. **8**, 227–235 (2014)
76. Ogryczak, W., Śliwiński, T.: On solving linear programs with the ordered weighted averaging objective. Eur. J. Oper. Res. **148**(1), 80–91 (2003)
77. Ogryczak, W., Śliwiński, T.: On efficient WOWA optimization for decision support under risk. Int. J. Approx. Reason. **50**, 915–928 (2009)
78. Papadimitriou, C.H., Steiglitz, K.: Combinatorial Optimization: Algorithms and Complexity. Dover Publications Inc., New York (1998)
79. Pereira, J., Averbakh, I.: Exact and heuristic algorithms for the interval data robust assignment problem. Comput. Oper. Res. **38**(1153–1163) (2011)
80. Pereira, J., Averbakh, I.: The robust set covering problem with interval data. Ann. Oper. Res. **207**, 217–235 (2013)
81. Perez-Galarce, F., Alvarez-Miranda, E., Candia-Vejar, A.: On exact solutions for the minmax regret spanning tree. Comput. Oper. Res. **47**, 114–122 (2014)
82. Şeref, O., Ahuja, R.K., Orlin, J.B.: Incremental network optimization: theory and algorithms. Oper. Res. **57**, 586–594 (2009)
83. Torra, V.: The weighted OWA operator. Int. J. Intell. Syst. **12**, 153–166 (1997)
84. Woeginger, G.J., Deineko, V.G.: On the robust assignment problem under a fixed number of cost scenarios. Oper. Res. Lett. **34**, 175–179 (2006)
85. Yager, R.R.: On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Trans. Syst. Man Cybern. **18**, 183–190 (1988)
86. Yaman, H., Karasan Oya, E., Pinar Mustafa, C.: The robust spanning tree problem with interval data. Oper. Res. Lett. **29**, 31–40 (2001)
87. Yu, G.: On the max-min 0–1 knapsack problem with robust optimization applications. Oper. Res. **44**, 407–415 (1996)
88. Yu, G., Kouvelis, P.: Complexity results for a class of min-max problems with robust optimization applications. In: Pardalos, P.M. (ed.) Complexity in Numerical Optimization. World Scientific, Singapore (1993)
89. Yu, G., Yang, J.: On the robust shortest path problem. Comput. Oper. Res. **6**, 457–468 (1998)
90. Zhu, G., Lue, X., Miao, Y.: Exact weight perfect matching of bipartite graph is NP-complete. In: Proceedings of the World Congress on Engineering, vol. 2, pp. 1–7 (2008)
91. Zieliński, P.: The computational complexity of the relative robust shortest path problem with interval data. Eur. J. Oper. Res. **158**, 570–576 (2004)

Chapter 7
Performance Analysis in Robust Optimization

André Chassein and Marc Goerigk

Abstract We discuss the problem of evaluating a robust solution. To this end, we first give a short primer on how to apply robustification approaches to uncertain optimization problems using the assignment problem and the knapsack problem as illustrative examples. As it is not immediately clear in practice which such robustness approach is suitable for the problem at hand, we present current approaches for evaluating and comparing robustness from the literature, and introduce the new concept of a scenario curve. Using the methods presented in this chapter, an easy guide is given to the decision maker to find, solve and compare the best robust optimization method for his purposes.

7.1 Introduction

Assume you have to solve a real-world optimization problem, which can be modeled, e.g., by an integer linear program. However, an optimal solution to this model might perform quite poorly in practice, as this first modeling approach neglects uncertainty in the problem parameters. Thus, some optimization tool that includes uncertain data is required. You have quite a range of methods to choose from: Stochastic optimization [12], fuzzy programming [14], interval programming [19], or robust optimization.

M. Goerigk (✉)
Lancaster University, Lancaster, UK
e-mail: m.goerigk@lancaster.ac.uk

A. Chassein
University of Kaiserslautern, Kaiserslautern, Germany
e-mail: chassein@mathematik.uni-kl.de

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_7

Assume you have chosen the last option. Again, there is a wide range of robust optimization concepts you may select: classic (strict) robustness [3], absolute or relative regret [22], adjustable robustness [6], recoverable robustness [23], light robustness [16], soft robustness [8], lexicographic α-robustness [20], recovery-to-optimality [17], or similarity-based robustness [13], to name some.

How to decide which of these approaches is appropriate for the problem at hand? Furthermore, most approaches are connected with some kind of "robust objective function" that decides on the degree of robustness. That is, every approach uses a different measure to decide which solution should be called robust, and which not. Also, many robust optimization approaches have parameters that control their "degree of robustness", but it is unclear in advance how to set them. To decide which one should actually be put into practice, you need some possibility to compare the robust solutions of each of these parameters and approaches.

In this chapter, we aim at shedding light on this problem of evaluating a robust solution from a bird's eye, or meta-perspective. We define a range of frameworks that allow to compare the performance of different robust solutions.

This work is not intended as a survey on robust optimization, for which we refer to [1, 7, 18] and [11]. Instead, having introduced some necessary notation in Sect. 7.2, we present a walk-through on the application of some of the most popular robust optimization approaches using an uncertain assignment problem as an example in Sect. 7.3. We discuss ways to evaluate a robust solution in Sect. 7.4, using an additional uncertain knapsack problem as an example. These evaluation frameworks are then illustrated using experimental data in Sect. 7.5.

7.2 Notations and Definitions

We first introduce the notation we use in this chapter to denote a general optimization problem. Additionally, we present two academic test problems we use to apply and to compare different robustness concepts. It is a common approach to use well studied academic test problems to compare new algorithms or concepts. On the one hand, these problems have an easy structure and are, therefore, easy to understand. On the other hand, it turns out that many real world problems are closely related variants of these problems.

7.2.1 General Notation

Consider the general minimization problem

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & g(x) \leq 0 \\ & x \in \mathscr{X}. \end{aligned}$$

In robust optimization it is assumed that the parameters ξ that describe the objective function or the constraints of the problem are not known exactly; instead, one assumes to know only a set $\mathscr{U}$ to which the parameters must belong to. These sets are called *uncertainty sets*.

We informally write the uncertain problem as

$$\begin{aligned} &\text{“min”} f(x,\xi) \\ &\text{s.t. “}g(x,\xi) \leq 0\text{”} \\ &x \in \mathscr{X}. \end{aligned}$$

As there exists no unique interpretation of the uncertain objective function and the uncertain constraints, it is not obvious how the robust problem should be solved. Different interpretations have been made, resulting in different robustness concepts. Some of these concepts are presented in Sect. 7.3.

The uncertainty set $\mathscr{U}$ can either be a finite list of different scenarios, which is denoted by *finite uncertainty* set, or a continuous sets. Continuous uncertainty sets are typically hyper-boxes, polytopes, or ellipsoids. A hyper-box uncertainty set is also called *interval uncertainty* set as it defines intervals for the different parameter values. We focus in this chapter mainly on finite and interval uncertainty sets. For the discussion of more advanced uncertainty sets we refer to [18].

We use the following notation for the uncertainty sets. Finite uncertainty sets are given by a list of scenarios $\mathscr{U}_F = \{\xi^1, \xi^2, \ldots, \xi^N\}$ and interval uncertainty sets are stated as $\mathscr{U}_I = \left\{\xi : \xi_j \in [\underline{\xi}_j, \overline{\xi}_j]\right\}$.

7.2.2 The Uncertain Assignment Problem

The assignment problem is defined by a complete bipartite graph with node sets V and W, $|V| = |W| = n$, and edge costs c_{ij} for all $i, j \in [n] := \{1, \ldots, n\}$. A feasible assignment is a subset of edges such that every node from V is connected to exactly one node from W (and vice versa). The problem is to find a feasible assignment that minimizes the sum of edge costs.

Written as an integer linear program (IP), the assignment problem can be stated as:

$$\begin{aligned} \text{(P)} \quad \min \ & \sum_{i\in[n]} \sum_{j\in[n]} c_{ij} x_{ij} && && (7.1) \\ \text{s.t.} \ & \sum_{i\in[n]} x_{ij} = 1 && \forall j \in [n] && (7.2) \\ & \sum_{j\in[n]} x_{ij} = 1 && \forall i \in [n] && (7.3) \\ & x_{ij} \in \{0,1\} && \forall i,j \in [n] && (7.4) \end{aligned}$$

Variable x_{ij} equals to 1 if and only if edge (i,j) is part of the assignment. Constraints (7.2) and (7.3) ensure that the assignment is feasible. That means that every element from V must be mapped to exactly one element of W, and vice versa. As the constraint matrix is totally unimodular, the integrality constraint (7.4) is equivalent to its relaxed version $x_{ij} \in [0,1]\ \forall i,j \in [n]$. The resulting problem is a linear program (LP). Thus, problem (P) can be solved in polynomial time.

In the uncertain optimization problem the exact knowledge of all edge costs c_{ij} is not given. Instead, we assume that these values are the result of some uncertain process. The set of all possible outcomes of c define the uncertainty set $\mathscr{U}$. As in Sect. 7.2.1, we use the following notation for finite and interval uncertainty sets: $\mathscr{U}_F = \{c^1,\dots,c^N\}$, and $\mathscr{U}_I = \times_{i,j\in[n]}[\underline{c}_{ij},\overline{c}_{ij}]$. The midpoint of $\mathscr{U}_I$ is denoted by $\hat{c} = 0.5(\underline{c}+\overline{c})$. We write P($c$) to denote the assignment problem with respect to the costs $c \in \mathscr{U}$.

Note that in this problem only the objective function is affected by uncertainty. If an assignment is feasible, it is feasible for all possible scenarios that might occur. This does not hold for the uncertain knapsack problem, which we explain in the following.

7.2.3 The Uncertain Knapsack Problem

The knapsack problem is defined by a set I of n items and a fixed budget B. Each item $i \in I$ is described by a positive weight w_i and a profit p_i. A packing is a subset of all items. A packing is feasible if the sum of weights of all items contained in this packing does not exceed the budget B. The problem is to find a feasible packing that maximizes the sum of all the profits, and can be stated as an integer program (IP):

$$\max \sum_{i\in[n]} p_i x_i \tag{7.5}$$

$$\text{s.t.} \sum_{i\in[n]} w_i x_i \leq B \tag{7.6}$$

$$x_i \in \{0,1\} \qquad \forall i \in [n] \tag{7.7}$$

Variable x_i equals to 1 if and only if item i is part of the packing. Constraint (7.6) ensures that the budget capacity B is not exceeded. Being NP-complete, the computational complexity of this problem is harder than for the assignment problem (for a general survey on the knapsack problem, see [21]).

In the uncertain version of this problem we assume that both item weights and item profits are affected by uncertainty. The uncertainty set $\mathscr{U}$ contains all possible combinations of weights and profits (p,w). We use the following notation for finite and interval scenarios sets: $\mathscr{U}_F = \{(p^1,w^1),(p^2,w^2),\dots(p^N,w^N)\}$ and $\mathscr{U}_I = (\times_i[\underline{p}_i,\overline{p}_i]) \times (\times_i[\underline{w}_i,\overline{w}_i])$.

Note that in this problem not only the objective function but also the constraints are affected by uncertainty. Hence, it is possible that packings are only feasible for some but not for all scenarios.

7.3 Approaches to Robust Optimization

In this section we present different robustness concepts that are compared in Sect. 7.4. The concrete solution of a robustness concepts needs the solution of a *robust counterpart*. The structure and the complexity of the robust counterpart depend greatly on the underlying uncertainty set that is used to describe the uncertainty. We use the assignment problem to illustrate the different concepts and the corresponding counterparts.

7.3.1 Strict Robustness

Also called *min-max* robustness or *classical* robustness, this is the most conservative way to solve an uncertain optimization problem (see [7]). This concepts asks for a solution that is feasible under all possible scenarios and gives the best performance guarantee, i.e. it optimizes the performance of the worst scenario for the chosen solution. This yields the following interpretation of the general robust optimization problem

$$\begin{aligned} &\min \max_{\xi \in \mathscr{U}} f(x, \xi) \\ &\text{s.t. } g(x, \xi) \leq 0 && \forall \xi \in \mathscr{U} \\ &\quad x \in \mathscr{X}. \end{aligned}$$

7.3.1.1 Finite Uncertainty

In the case of a finite uncertainty set $\mathscr{U}_F$, the general optimization problem attains the form

$$\begin{aligned} &\min \max_{k \in [N]} f(x, \xi^k) \\ &\text{s.t. } g(x, \xi^k) \leq 0 && \forall k \in [N] \\ &\quad x \in \mathscr{X}. \end{aligned}$$

It turns out that this uncertainty can lead to very difficult robust counterparts. Several negative complexity results are shown even if $\mathscr{U}$ consists of only two scenarios and the underlying problems are very basic [1]. Nevertheless, the robust counterpart can

be formulated in most cases as a mixed integer programming (MIP) problem. We showcase this in the following for the assignment problem.

$$\min z \tag{7.8}$$

$$\text{s.t.} \sum_{i\in[n]}\sum_{j\in[n]} c_{ij}^k x_{ij} \le z \qquad \forall k \in [N] \tag{7.9}$$

$$\sum_{i\in[n]} x_{ij} = 1 \qquad \forall j \in [n] \tag{7.10}$$

$$\sum_{j\in[n]} x_{ij} = 1 \qquad \forall i \in [n] \tag{7.11}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \tag{7.12}$$

$$z \ge 0 \tag{7.13}$$

Constraint (7.9) ensures that variable z is equal to the worst performance of solution x for all possible scenarios $c^1,\dots,c^N$ in an optimal solution. This additional constraints destroys the total unimodularity of the constraint matrix. Hence, one cannot relax the integrality constraints anymore; in fact, one can show that this problem is NP-complete already for two scenarios [22].

7.3.1.2 Interval Uncertainty

In the case of interval uncertainty $\mathcal{U}_I$ the robust counterpart can be stated as a semi-infinite program, which is an optimization problem with finitely many variables and infinitely many constraints.

$$\begin{aligned} \min\ & z \\ \text{s.t. } & f(x,\xi) \le z && \forall \xi \in \mathcal{U}_I \\ & g(x,\xi) \le 0 && \forall \xi \in \mathcal{U}_I \\ & x \in \mathcal{X}. \end{aligned}$$

In special cases this problem can be greatly simplified. If x is always positive and the objective function has the form $f(x,\xi) = \xi^t x$, the infinitely many constraints describing the objective function can be replaced by one. It suffices to consider $f(x,\overline{\xi}) \le z$, as the worst scenario that might happen for any solution is scenario $\overline{\xi}$.

The presented assignment problem fulfills these properties. In the worst case scenario the edge costs are given by $\overline{c}$. Hence, the robust counterpart reduces to the problem P($\overline{c}$). As this problem has the structure of the original, certain problem, it can be solved with the same algorithms in polynomial time.

The following two approaches rely on the idea to reduce the size of the uncertainty sets. An illustration of both approaches is given in Fig. 7.1.

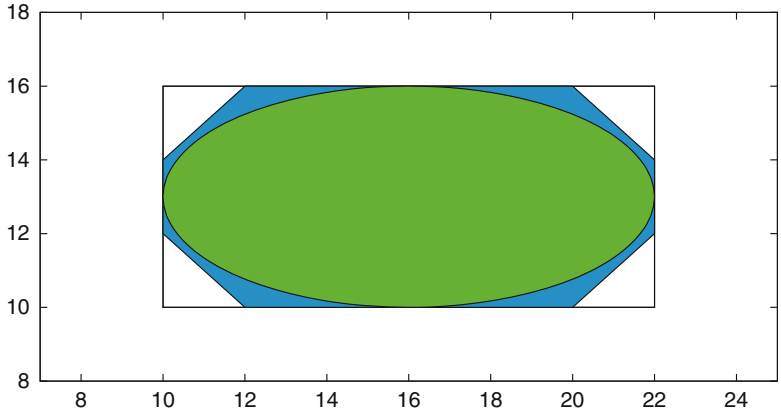


Fig. 7.1: Cutting unlikely corners. The *rectangle* represents the complete interval uncertainty set. The bounded uncertainty (see Sect. 7.3.2) set is shown as the *blue* polytope, the ellipsoidal uncertainty set (see Sect. 7.3.3) is represented by the *green* ellipsoid

7.3.2 Bounded Uncertainty

This approach was introduced by Bertsimas and Sim [10]. They motivate their approach with the observation that the strict robustness concept with interval uncertainty is very pessimistic, as it assumes that all parameters attain their worst possible value at the same time. As this seems to be an unrealistic assumption for many real-world situations, they suggest to introduce another uncertainty set that bounds the deviation of the parameters.

We present this idea using the assignment problem. For integral values of Γ, the resulting uncertainty set for the assignment problem has the following form:

$$\mathscr{U}_I(\Gamma) = \left\{ c \in \mathscr{U}_I : \left| \left\{ (i,j) : c_{ij} > \hat{c}_{ij} \right\} \right| \leq \Gamma \right\}$$

i.e., the number of coefficients that are larger than in the midpoint scenario is bounded by the parameter Γ. The concept can also be generalized to non-integral Γ values.

Using this uncertainty set the robust counterpart of the assignment problem is given by the following mixed integer program (MIP).

$$\min \sum_{i\in[n]} \sum_{j\in[n]} \hat{c}_{ij} x_{ij} + \Gamma\pi + \sum_{i\in[n]} \rho_{ij} \tag{7.14}$$

$$\text{s.t.} \sum_{i\in[n]} x_{ij} = 1 \qquad \forall j \in [n] \tag{7.15}$$

$$\sum_{j\in[n]} x_{ij} = 1 \qquad \forall i \in [n] \tag{7.16}$$

$$\pi + \rho_{ij} \geq (\overline{c}_{ij} - \hat{c}_{ij})x_{ij} \qquad \forall i,j \in [n] \qquad (7.17)$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \qquad (7.18)$$

$$\rho_i \geq 0 \qquad \forall i \in [n] \qquad (7.19)$$

$$\pi \geq 0 \qquad (7.20)$$

Additional to x, variables π and ρ are introduced. If Γ is chosen as large as the number of x-variables (i.e., $\Gamma = n^2$), then π is equal to 0 in an optimal solution. In this case ρ_{ij} is equal to $(\overline{c}_{ij} - \hat{c}_{ij})x_{ij}$, which is guaranteed by Inequality (7.17). Replacing ρ_{ij} accordingly in the objective function, only $\overline{c}^t x$ remains, i.e. the worst case objective function. This was expected as for such a large value of Γ the bounded uncertainty set is equal to the original uncertainty set, hence, the model reduces to the strict robust counterpart.

Contrary, if Γ is set to 0, π can be made arbitrary large in an optimal solution. Inequality (7.17) is superfluous and ρ_{ij} can be set to 0. The objective function reduces to $\hat{c}^t x$. Also this was expected, as for Γ equal to 0, only $\hat{c}^t x$ is contained in the bounded uncertainty set. In [9, 10] it is explained in detail how this formulation can be derived. There, it is also shown that the resulting robust problem is solvable in polynomial time. Note that this concept is also applicable if constraints are affected by uncertainty.

7.3.3 Ellipsoidal Uncertainty

The use of ellipsoidal uncertainty sets can be motivated from two different reasons. The first one is that many uncertainty sets are already ellipsoidal in practice, e.g., when stemming from a normal data distribution. The second one follows the idea of bounded uncertainty. Even if you have given an interval uncertainty set it can be a good idea to use an ellipsoidal uncertainty set to cut off unlikely corners.

For more information about ellipsoidal uncertainty sets we refer to the papers of Ben-Tal and Nemirovski [4, 5]. We use again the assignment problem to present the resulting robust counter part if an ellipsoidal uncertainty set is used to cut off unlikely corners of the interval uncertainty set $\mathcal{U}_I$. As before, the midpoint of $\mathcal{U}_I$ is denoted by $\hat{c}$.

$$\min \sum_{i\in[n]} \sum_{j\in[n]} \overline{c}_{ij}x_{ij} - \sum_{i,j\in[n]} (\overline{c}_{ij} - \hat{c}_{ij})p_{ij} + \Omega q \qquad (7.21)$$

$$\text{s.t.} \sum_{i,j\in[n]} (\overline{c}_{ij} - \hat{c}_{ij})^2 p_{ij}^2 \leq q^2 \qquad (7.22)$$

$$0 \leq p_{ij} \leq x_{ij} \qquad \forall i,j \in [n] \qquad (7.23)$$

$$0 \leq q \qquad (7.24)$$

$$\sum_{i\in[n]} x_{ij} = 1 \qquad \forall j \in [n] \tag{7.25}$$

$$\sum_{j\in[n]} x_{ij} = 1 \qquad \forall i \in [n] \tag{7.26}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \tag{7.27}$$

Ben-Tal and Nemirovski explain in [4] how to derive this problem formulation. The parameter Ω controls the size of the ellipsoid that is used to approximate $\mathscr{U}_I$. If Ω is set to 0 the ellipsoid consists of the single point $\hat{c}$. In this case the robust counterpart reduces to $P(\hat{c})$. On the other hand, if Ω is large enough, the problem becomes the strict robust problem $P(\overline{c})$. Note that the problem formulation contains a quadratic constraint. Therefore, it can not be solved anymore using mixed integer programming. Nevertheless, if the integer constraints are relaxed one obtains a convex program that can be solved efficiently.

7.3.4 Regret Robustness

To apply the regret robustness concept it is assumed that only the objective function is affected by uncertainty and the constraints are certain. In strict robustness the evaluation of a solution depends solely on the performance under one special scenario. It is neglected that this special scenario might also be bad for all other possible solutions. Hence, it could be meaningful to take into account the best possible performance that could be achieved for this special scenario. This idea is used in regret robustness. For a fixed scenario, the regret of a solution is computed using both the objective function of the solution and the best possible objective value. There exist different methods to compute the regret of a solution. We present three in the following.

7.3.4.1 Absolute Regret

In absolute regret robustness, one adds some normalization to the robust objective value, so that taking the maximum over all scenarios becomes "more fair". Specifically, we consider the robust objective function

$$reg(x) = \max_{\xi\in\mathscr{U}} f(x,\xi) - opt(\xi)$$

where $opt(\xi)$ denotes the best possible objective value for the problem that is described by parameter ξ.

This objective function yields to a different interpretation of the general robust optimization problem.

$$\begin{aligned} &\min \max_{\xi \in \mathscr{U}} f(x,\xi) - opt(\xi) \\ &\text{s.t. } g(x) \leq 0 \\ &\quad x \in \mathscr{X}. \end{aligned}$$

Finite and interval uncertainty sets lead again to different robust counterparts. We present this with the assignment problem.

For a finite set of scenarios $\mathscr{U}_F$ of polynomial size, the optimal objective values can be precomputed in polynomial time. The resulting robust counterpart is very similar to the strict robust counterpart:

$$\min z \tag{7.28}$$

$$\text{s.t. } \sum_{i\in[n]}\sum_{j\in[n]} c_{ij}^k x_{ij} - opt(c^k) \leq z \qquad \forall k \in [N] \tag{7.29}$$

$$\sum_{i\in[n]} x_{ij} = 1 \qquad \forall j \in [n] \tag{7.30}$$

$$\sum_{j\in[n]} x_{ij} = 1 \qquad \forall i \in [n] \tag{7.31}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \tag{7.32}$$

$$z \geq 0 \tag{7.33}$$

Again, variable z is introduced to monitor the robust objective function. Constraint (7.29) ensures that z equals to the maximum regret in an optimal solution. Specialized algorithms such as branch and bound algorithms can be applied to this problem formulation. However, this is quite different for interval-uncertainty sets, where it is not possible to compute all values $opt(c)$ in advance. Nevertheless, it is possible to formulate the resulting robust counterpart as:

$$\min \sum_{i\in[n]}\sum_{j\in[n]} \overline{c}_{ij} x_{ij} - \sum_{i\in[n]} (\alpha_i + \beta_i) \tag{7.34}$$

$$\text{s.t. } \alpha_i + \beta_j \leq \underline{c}_{ij} + (\overline{c}_{ij} - \underline{c}_{ij}) x_{ij} \qquad \forall i,j \in [n] \tag{7.35}$$

$$\sum_{i\in[n]} x_{ij} = 1 \qquad \forall j \in [n] \tag{7.36}$$

$$\sum_{j\in[n]} x_{ij} = 1 \qquad \forall i \in [n] \tag{7.37}$$

$$\alpha_i, \beta_i \geqslant 0 \qquad \forall i \in [n] \tag{7.38}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \tag{7.39}$$

where x, α, and β are variables. To derive this problem formulation it is used that the scenario that maximizes the regret of a solution is described by the following rule: All elements that are chosen by the solution x are as expensive as possible, and all other elements as cheap as possible. The recipe that was used to derive this MIP formulation can be applied to any combinatorial optimization problem with

uncertainty in the costs, and for which the nominal problem (P) can be solved by using its linear relaxation. For more information about the derivation of the robust counterpart that arises from the absolute regret concept we refer to [25].

7.3.4.2 Relative Regret

The previous absolute regret approach aims at a normalization of objective values by using the difference to the best possible objective value in any scenario. However, this normalization may not be appropriate for some applications. The relative regret normalizes the absolute regret by dividing it with the best possible objective function under a scenario. The relative regret objective function has the following form for a general optimization problem.

$$rreg(x) = \max_{\xi \in \mathscr{U}} \frac{f(x,\xi) - opt(\xi)}{opt(\xi)}$$

The relative regret concept breaks down to the absolute regret concept for finite uncertainty sets with a different scaling of scenarios. But for interval uncertainty sets this does not hold. For this case, we present a possible formulation of the relative regret robust counterpart for the assignment problem.

$$\min \mu \tag{7.40}$$

$$\text{s.t.} \sum_{i\in[n]} \sum_{j\in[n]} \overline{c}_{ij} x_{ij} \leq \sum_{i\in[n]} (\alpha_i + \beta_i) \tag{7.41}$$

$$\alpha_i + \beta_j \leq \mu \underline{c}_{ij} + (\overline{c}_{ij} - \underline{c}_{ij}) x_{ij} \qquad \forall i,j \in [n] \tag{7.42}$$

$$\sum_{i\in[n]} x_{ij} = 1 \qquad \forall j \in [n] \tag{7.43}$$

$$\sum_{j\in[n]} x_{ij} = 1 \qquad \forall i \in [n] \tag{7.44}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \tag{7.45}$$

$$\alpha, \beta, \mu \gtrless 0 \tag{7.46}$$

Deriving this formulation of the problem is more involved than in the case of absolute regret, but solving it seems to be almost of the same computational complexity. An additional variable μ is introduced that represents the ratio of objective function and optimal objective value. In [2] one can find the detailed derivation of this problem formulation.

7.3.4.3 Alpha Regret

For discrete uncertainty sets there exist another approach to interpret how the regret of a solution should be calculated. The alpha regret concept is similar to

the concept of absolute regret, but extends it by the notion of anonymization. The idea is to compare the realized solution not with the optimal solution that could be realized in the same scenario. Instead, the vector of solution values $V(x) = \big(f(x,\xi^1), f(x,\xi^2), \ldots, f(x,\xi^N)\big)$ and the vector of optimal solution values $V^* = \big(opt(\xi^1), opt(\xi^2), \ldots, opt(\xi^N)\big)$ are both sorted and then compared in each component. The maximum difference is called the alpha regret. By comparing the solution of the kth best scenario with the kth best optimal solution, the scenarios are made anonymous. This is plausible if it is not known in advance, which scenario will happen or is more likely to happen. Formally, the alpha regret of a solution can be computed as

$$\alpha reg(x) = \min_{\pi \in \sigma(N)} \max_{i \in [N]} f(x,\xi^i) - opt(\xi^{\pi(i)}),$$

where $\sigma(N)$ denotes the set of all permutations of the set $[N]$. Hence, the resulting formulation for the general optimization problem looks as follows

$$\begin{aligned} &\min z \\ &\text{s.t. } f(x,\xi^i) - f^*(\xi^{\pi(i)}) \le z && \forall i \in [N] \\ &\quad g(x) \le 0 \\ &\quad x \in \mathscr{X} \\ &\quad \pi \in \sigma(\mathscr{U}) \end{aligned}$$

To give a concrete example we use again the assignment problem.

$$\min z \tag{7.47}$$

$$\text{s.t. } \sum_{i\in[n]} \sum_{j\in[n]} c^\ell_{ij} x_{ij} - \sum_{k=1}^{N} p_{k\ell}\, opt(c^k) \le z \qquad \forall \ell \in [N] \tag{7.48}$$

$$\sum_{k\in[N]} p_{k\ell} = 1 \qquad \forall \ell \in [N] \tag{7.49}$$

$$\sum_{\ell\in[N]} p_{k\ell} = 1 \qquad \forall k \in [N] \tag{7.50}$$

$$\sum_{i\in[n]} x_{ij} = 1 \qquad \forall j \in [n] \tag{7.51}$$

$$\sum_{j\in[n]} x_{ij} = 1 \qquad \forall i \in [n] \tag{7.52}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \tag{7.53}$$

$$p_{k\ell} \in \{0,1\} \qquad \forall k,\ell \in [N] \tag{7.54}$$

The variables $p_{k\ell}$ are used to represent the possible permutations of the scenarios. Variable $p_{k\ell}$ is set to 1 if the kth scenario is sorted to position ℓ, i.e. if $\pi(k) = \ell$. The alpha regret concept is introduced in [20].

7.3.5 Recoverable Robustness

In the previous approaches, we are interested in finding a single solution, that is supposed to perform well under all possible scenario outcomes. It is not possible to modify this solution, once the actual scenario becomes known. In two-stage approaches to robust optimization (see also the approach of adjustable robustness), this possibility is included in the model. Once the scenario is revealed, we can do some modifications to our solutions. Naturally, if we could change the complete solution, we could simply recover to an optimal solution in any scenario. Thus, the amount of modifications that we can perform is usually bounded. Typically, one considers the min-max objective over all scenarios in this setting; however, any other objective function such as absolute or relative regret would also be conceivable. For the general uncertain optimization problem this yields an infinite program with infinitely many variables and constraints.

$$
\begin{aligned}
&\min z \\
&\text{s.t. } f(x_\xi, \xi) \le z && \forall \xi \in \mathscr{U} \\
&\quad g(x_\xi, \xi) \le 0 && \forall \xi \in \mathscr{U} \\
&\quad dist(x, x_\xi) \le D && \forall \xi \in \mathscr{U} \\
&\quad x_\xi \in \mathscr{X} \\
&\quad x \in \mathscr{X}
\end{aligned}
$$

The solution to the recovery robust problem is given by x. For each possible scenario ξ a variable x_ξ is introduced. Each solution x_ξ must itself be feasible and close to the solution x. The function $dist$ is used to measure how close two solutions are. The maximum allowed distance is given by parameter D.

For the assignment problem with finite uncertainty sets the problem can be written as a MIP. We assume that we are allowed to modify up to $2K$ variables x_{ij} once the scenario is known, i.e. we can remove K choices, and add K new choices to our solution. The resulting problem is given as

$$
\begin{aligned}
&\min z && && (7.55)\\
&\text{s.t. } \sum_{i\in[n]} \sum_{j\in[n]} c_{ij}^k x_{ij}^k \le z && \forall k \in [N] && (7.56)\\
&\quad \sum_{i\in[n]} x_{ij} = 1 && \forall j \in [n] && (7.57)\\
&\quad \sum_{j\in[n]} x_{ij} = 1 && \forall i \in [n] && (7.58)\\
&\quad \sum_{i\in[n]} x_{ij}^k = 1 && \forall j \in [n], k \in [N] && (7.59)\\
&\quad \sum_{j\in[n]} x_{ij}^k = 1 && \forall i \in [n], k \in [N] && (7.60)
\end{aligned}
$$

$$-y_{ij}^k \leq x_{ij} - x_{ij}^k \leq y_{ij}^k \qquad \forall i,j \in [n], k \in [N] \tag{7.61}$$

$$\sum_{i\in[n]} \sum_{j\in[n]} y_{ij}^k \leq 2K \qquad \forall k \in [N] \tag{7.62}$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in [n] \tag{7.63}$$

$$x_{ij}^k \in \{0,1\} \qquad \forall i,j \in [n], k \in [N] \tag{7.64}$$

$$y_{ij}^k \in \{0,1\} \qquad \forall i,j \in [n], k \in [N] \tag{7.65}$$

We use variables x to model the first-stage solution, and variables x^k for every scenario $k \in [N]$, to model the second-stage (adapted) solutions. The auxiliary variables y^k are used to measure the difference between x and x^k. In Constraints (7.57) and (7.58), we ensure that our first-stage solution x is a feasible assignment, while Constraints (7.59) and (7.60) ensure the same for each scenario. Constraints (7.61) and (7.62) bound the difference between first- and second-stage solutions. More about recovery robust optimization can be found in [23].

7.3.6 Summary

We discussed numerous different concepts for robust optimization—still, the presented list of concepts is not exhaustive. Other interesting concepts can be found for example in [18], and in Sect. 7.1 of this chapter.

In this section we provide a short overview about all presented concepts. In Table 7.1 we highlight under which uncertainty contexts the different concepts are applicable.

	Cons & Obj		Obj	
	$\mathcal{U}_F$	$\mathcal{U}_I$	$\mathcal{U}_F$	$\mathcal{U}_I$
Strict robustness	✓	✓	✓	✓
Bounded uncertainty	–	✓	–	✓
Ellipsoid uncertainty	a)	✓	a)	✓
Regret robustness				
Absolute regret	–	–	✓	✓
Relative regret	–	–	b)	✓
Alpha regret	–	–	✓	–
Recoverable robustness	✓	✓	✓	✓

Table 7.1: This table shows under which uncertainty context the different robustness concepts are applicable. The columns with the label Cons & Obj refer to the case where both the constraints and the objective function are affected by uncertainty. The columns with the label Obj refer to the case where only the objective function is affected by uncertainty. a) An ellipsoid can be computed that contains all point from the discrete set. This will guarantee a safe approximation of the original problem. b) The concept of relative regret reduces to the concept of absolute regret for finite uncertainty sets

7.4 Frameworks to Evaluate Robust Solutions

Every robustness concept is motivated from a different perspective and has its own benefits and drawbacks. Therefore, it is unclear how these different concepts perform in comparison to each other. To make two concepts comparable one has to define a framework in which the quality of the solutions that are produced by the different concepts can be measured. In this section we give a short introduction to some of these frameworks. We discuss them in more detail in the experimental Sect. 7.5.

Two robustness concepts can only be compared if the used uncertainty context is applicable for both (see Table 7.1). Hence we define different kind of frameworks for different uncertainty contexts.

For some frameworks we assume the knowledge of an average case scenario, also called *nominal* scenario. The performance of the solution under the nominal scenario is an important indication for the overall quality of the solution.

If we want to speak about feasibility probability we have to make assumptions on the underlying probability structure of the problem.

What is assumed to be known for the different frameworks is given in Table 7.2.

Framework	Nominal scenario	Probability distribution
Price of robustness	✓	✓
AC-WC curve	✓	–
Scenario curve	–	–
Sampled scenario curve	–	✓
Scenario curve with recovery	✓	–

Table 7.2: Assumptions for the different frameworks

7.4.1 The Price of Robustness

This framework is applicable if constraints are affected by uncertainty. Stemming from the seminal paper carrying the same name [10], the price of robustness (PoR) is defined as "the tradeoff between the probability of violation and the effect to the objective function of the nominal problem". This idea can be used to measure the quality of a solution. First, all solutions that have to be compared are evaluated with respect to their nominal performance. Next, simulation can be used to compute the probability of violation. The solutions are compared by drawing them into a two dimensional coordinate system. The y-axis defines the nominal performance of a solution and the x-axis gives the violation probability.

We note that the term "price of robustness" is also used differently, see, e.g., [24].

7.4.2 The AC-WC Curve

This framework is only applicable if the objective function is affected by uncertainty. If constraints are not affected by uncertainty, solutions are feasible for all scenarios; hence, it is not meaningful any more to speak about probability of violation. Instead, one can use the performance in the worst case to compare solutions. Solutions that need to be compared are drawn into a two dimensional coordinate system. The x-axis gives the nominal performance guarantee of a solution and the y-axis gives the performance in the worst case scenario.

We call a solution non-dominated if there exist no other solution that has both a better average and worst case performance. The set of all non-dominated solutions is defined as the *AC-WC curve*. The AC-WC curve can be computed effectively if the feasibility set of the problem is convex and some further technical assumptions are fulfilled. For more information about the AC-WC curve we refer to [15].

7.4.3 The Scenario Curve

For this framework it is assumed that the objective function is affected by uncertainty and the uncertainty set is finite. As the uncertainty set is finite the performance of a solution can be evaluated for each possible scenario to obtain the vector $F(x) = \left(f(x,\xi^1),\ldots,f(x,\xi^N)\right)$. Next we use the idea of anonymization: Similar to the concept of alpha regret, the vector $F(x)$ is sorted from good to bad performance. The sorted version of $F(x)$ is denoted by $F_s(x)$. To compare two different solutions x and x' the vectors $F_s(x)$ and $F_s(x')$ are drawn into a two dimensional coordinate system. The k^{th} component of vector $F_s(x)$ is represented by the point $(k,(F_s(x))_k)$. The leftmost point corresponds to the performance under the best scenario and the rightmost point corresponds to the performance under the worst scenario. The sorting of the different solutions leads to a better visualization of the solution quality.

Additionally, the vector $F^* = \left(opt(\xi^1),\ldots,opt(\xi^N)\right)$ can be computed and the sorted version of this vector is drawn into the same plot. This creates an optimal benchmark curve, that can be used for comparison.

7.4.4 The Sampled Scenario Curve

Using a sampling procedure the concept of the scenario curve can be transferred to arbitrary uncertainty sets. One needs to be able to sample a set $S = \left\{\tilde{\xi}^1,\ldots,\tilde{\xi}^K\right\}$ of possible parameter realizations. The sampled scenarios are then used to draw the sampled scenario curve.

If a solution is infeasible for a certain scenario, a "bad" value is assigned to this solution. This value must be considerably worse than the worst value of all feasible solutions (e.g., a profit of 0 for the knapsack problem).

7.4.5 The Scenario Curve with Recovery

We now discuss a second extension of the scenario curve approach from Sect. 7.4.3 to optimization problems with uncertainty in the constraints. To this end, we use the uncertain knapsack problem as an illustrative example again.

In these circumstances, it may happen that the robust solution we would like to evaluate is not feasible for some scenarios. We therefore assume that a recovery action is available: By changing up to K many items, we can manipulate our solution for every scenario. For every such recovery distance K, we can calculate a scenario curve as before, which results in a 2-dimensional scenario curve overall.

More precisely, we suggest the following approach to evaluate a solution x. We calculate optimal objective values for every scenario in the uncertainty set and sort these values. For every possible recovery distance $K = 1, \ldots, K_{max}$ we do the following: We calculate the best possible objective value of x for every scenario after the recovery action. Next we sort these values and normalize them using the sorted vector of optimal solutions. In this way we generate K_{max} scenario curves for solution x. We plot all these curves in one plot using a heat map.

On the horizontal axis is the recovery distance, and on the vertical axis are the sorted scenarios. Bright colors mean that the solution is close to the optimal solution after the recovery. A black field means that the solution could not be recovered to a feasible solution in this scenario for the given recovery budget.

7.5 Experiments

In this last section we use the uncertain assignment problem and the uncertain knapsack problem to illustrate the different frameworks. We use two different uncertainty setups for the assignment problem and four for the knapsack problem. We consider finite and interval uncertainty for the assignment problem and for the knapsack problem. For the knapsack problem we consider either profit uncertainty or profit and weight uncertainty.

Please keep in mind that all figures show the performance of the different concepts only for specific instances. Therefore, we avoid to make general statements about the different concepts. Instead, we want to explain how the different frameworks can be used to chose the best solution for the specific instances. Beside the different robustness concepts, we also compute the naive solution of the nominal scenario for comparison. This solution is called the average case solution.

7.5.1 Assignment Problem

We use an assignment instance with 40 nodes, 20 nodes on each side. For the finite uncertainty set we sample 10 different costs for each edge. The costs of an edge are chosen uniformly at random from the interval $[50, 150]$ for each scenario. The nominal scenario is chosen to be the average of the 10 sampled scenarios. In the case of interval uncertainty the midpoint of each interval is chosen uniformly at random from the interval $[100, 150]$ and the length of the interval is also chosen randomly such that edges that are cheap on average tend to have longer intervals.

7.5.1.1 Finite Uncertainty

We apply the following robustness concepts for this setting: Strict robustness, absolute/relative/alpha regret, and recoverable robustness. The recovery budget for recoverable robustness was set to 2. We present the scenarios curve in Fig. 7.2.

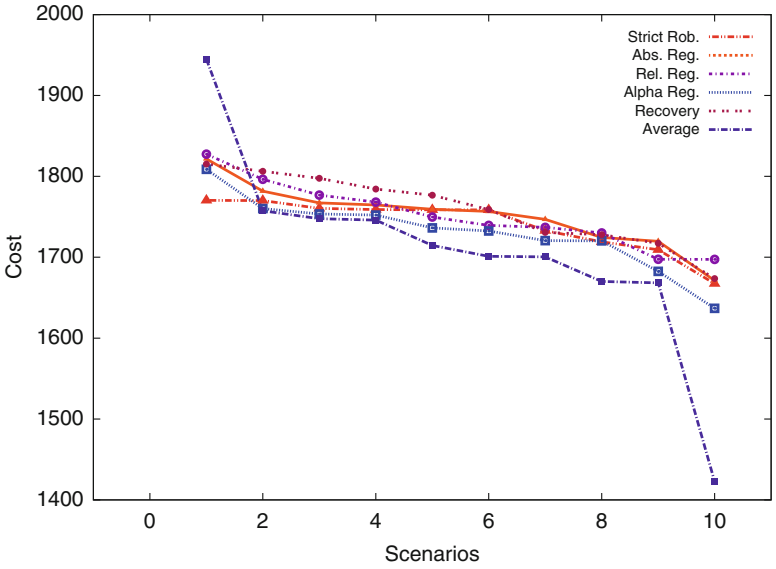


Fig. 7.2: Scenario curve of an assignment instance with 40 nodes an 10 cost scenarios

On the left side, the worst scenario for all solutions is compared, and on the right side, the best one (as this is a minimization problem, smaller values indicate better performance). As expected, strict robustness generates the solution that performs at best if the worst may happen. The other robustness concepts perform relatively similar in their worst scenario.

Let us compare the performance of the strict robust solution and the average case solution in more detail to highlight some aspects of robust optimization. The performance of the strict solution shows only little deviation among the different scenarios, whereas the average case solution has the largest performance deviation of all compared solutions. The average case solution is better in all but one scenario compared to the strict robust solution. If one is willing to accept the risk of bad performance in few scenarios, the average case solution may be an appropriate choice. But if this risk cannot be taken, one should rely on a robust solution.

For this instance, an interesting alternative to the strict robust solution is given by the alpha regret solution. It performs better in all but the worst scenario. Further, the performance in the worst scenario is still relatively close to the performance of the strict robust solution.

7.5.1.2 Interval Uncertainty

We apply the following robustness concepts for this setting: Strict robustness, bounded uncertainty, ellipsoidal uncertainty, absolute and relative regret. The parameter Γ describing the bounded uncertainty concept is chosen from the set $\{1,2,\ldots,10\}$. The parameter Ω defining the size of the ellipsoid used in the ellipsoidal uncertainty concept is chosen from the set $\{0.5,1.0,1.5,2.0\}$. Figure 7.3 shows the AC-WC curve of the instance.

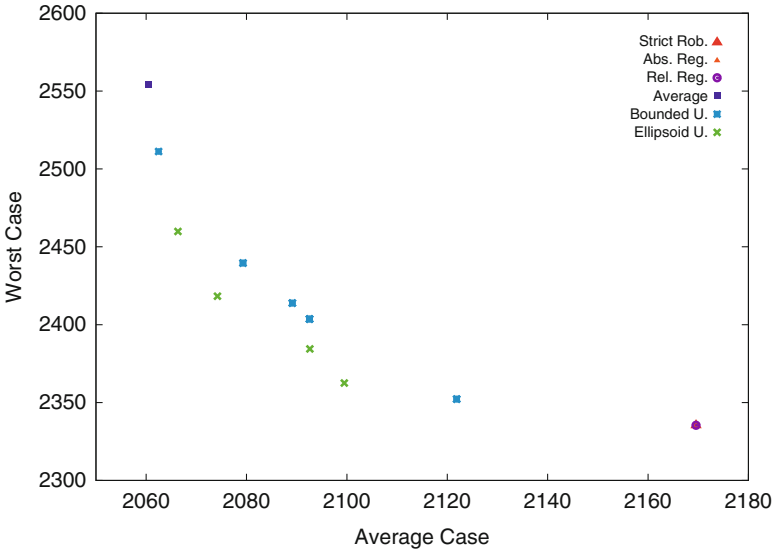


Fig. 7.3: The AC-WC curve of an assignment instance with 40 nodes and interval uncertainty. The solution of the Strict Robustness, Absolute and Relative regret concept coincide

For this instance strict robustness, absolute and relative regret generate the same solution. Nevertheless, it is interesting to compare the solutions of bounded and ellipsoidal uncertainty for different parameter choices. Small values of Γ resp. Ω produce solutions that are closer to the average case solution and larger values lead to solutions close to the strict solution. The AC-WC curve enables us to visualize the exact trade-off for moving from average to worst case optimization. For this instance, the bounded uncertainty concept generates solutions that are often dominated by solutions from the ellipsoidal uncertainty concept.

7.5.2 Knapsack Problem

The capacity of the knapsack is set to 500 in all instances. The following setup is used for finite uncertainty. If both weights and profits are affected by uncertainty, we use an instance with 50 items. For each scenario, the profit of an item is chosen uniformly at random from the interval $[50, 150]$ and the weight from the interval $[15, 25]$. If only the profits are affected by uncertainty, we use an instance with 200 items, where the weight of each item is chosen uniformly at random from the interval $[15, 25]$. The profits are generated as before. In both cases, we sample 10 scenarios.

For interval uncertainty we use 500 items. In this case we sample the average profit of an item uniformly from the interval $[60, 140]$ and the average weight from the interval $[12, 28]$. The length of the intervals is chosen randomly proportional to the midpoint of the interval. If the weights are not affected by uncertainty, the interval length is set to 0.

7.5.2.1 Finite Profit Uncertainty

We apply the following robustness concepts for this setting: Strict robustness, absolute/relative/alpha regret, and recoverable robustness. The recovery budget is set to 2. The scenario curve is shown in Fig. 7.4.

Again the performance in the worst scenario is shown on the left and the performance in the best on the right (as this is a maximization problem, larger values indicate better performance). As expected, the strict robust solution is the best in the worst case. Interestingly, this is not only true for the worst but also for the three worst scenarios, for this instance. Again the average case solution shows the largest performance deviation of all solutions. If we compare the absolute and the relative regret solution, none of them is clearly preferable. The absolute regret solution is preferable for good scenarios and the relative regret solution is preferable for bad scenarios. It is interesting to note that the strict robust solution performs unexpectedly well for its best scenario.

7.5.2.2 Finite Uncertainty

We apply the following robustness concepts for this setting: Strict robustness and recoverable robustness. The recovery budget is set to 10. The scenario curves with

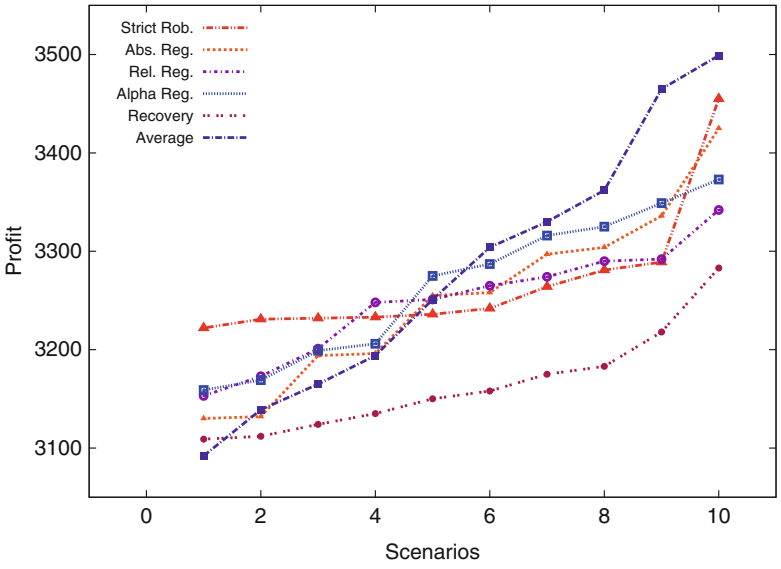


Fig. 7.4: The scenario curve of an knapsack problem with 200 items and 10 profit scenarios

recovery are shown in Fig. 7.5 for both solutions. Values are normalized with respect to the optimal benchmark curve, where yellow indicates good performance, and darker colors indicate worse performance.

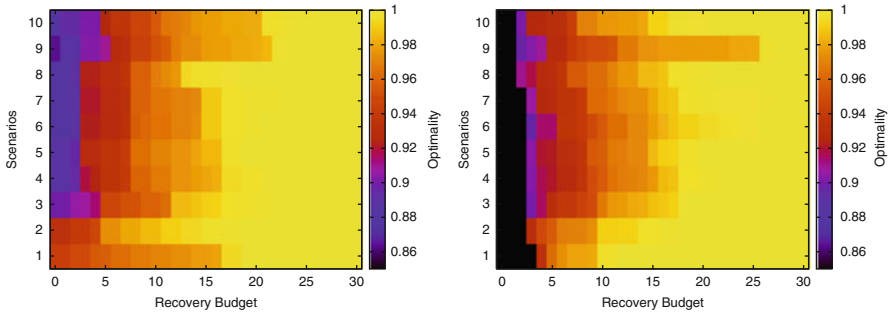


Fig. 7.5: The scenario curve with recovery of an knapsack problem with 50 items and 10 profit and weight scenarios. The *left figure* represents the strict robust solution, the *right figure* the solution generated by the recovery robust concept

The strict robust solution is ensured to be feasible for all scenarios. Hence, no black boxes can occur in the left figure. The recovery robust solution only guarantees feasibility within the recovery budget that was used for the computation. In this case we used a recovery budget of 10. This means that black boxes may occur in columns 0–9, which is indeed the case on the right figure. Note that the recovery robust solution is not feasible for a single scenario if recovery is not allowed. But, if a recovery budget of 10 is allowed, the recovery robust solution performs considerably better as the worst case solution in the worst case. This can be seen by comparing row 1 of both figures. If the allowed recovery budget is large enough, the originally chosen solution becomes irrelevant as recovery to the optimal solution is possible for each scenario. This explains the bright right side of both figures. Observe that the strict robust solution optimizes purely the worst case without recovery. Therefore this solution has the best value in the field corresponding to scenario 1 and a recovery budget of 0 compared to all other solutions. Whereas the recovery robust solution optimizes the worst case performance if a recovery budget of 10 is allowed. Hence, it has the best value in the field corresponding to scenario 1 and a recovery budget of 10 in comparison with all other solutions.

7.5.2.3 Interval Profit Uncertainty

We apply the following robustness concepts for this setting: Strict robustness, bounded uncertainty and ellipsoidal uncertainty. We chose $\Gamma = 15$ and $\Omega = 4$. We use this instance to present the sampled scenario curve. We sampled 1000 scenarios. Remember that the profits of the items are defined by intervals. In each scenario that we sample the profit of an item is equally likely one of the endpoints of the interval. The sampled scenario curve is shown in Fig. 7.6.

We also show the "optimal curve" that can be used for comparison. We separately solve each of the 1000 sampled scenarios. The resulting vector of performances is sorted and plotted in the figure. It is clear that no solution can generate a point that lies above the optimal curve. The sampled scenario curve visualizes the conservatism of the strict robust solution for interval uncertainty. For all sampled scenarios the solutions from the bounded and the ellipsoidal uncertainty concept perform better. Only the average case solution performs worse for some scenarios. The sampled scenario curve shows clearly the benefit of the bounded and ellipsoidal uncertainty approach, as very unlikely scenarios that will never happen in practice are ignored. It is interesting to note how similar the solutions generated by the bounded and ellipsoid uncertainty concept are, if the describing parameters are chosen accordingly.

7.5.2.4 Interval Uncertainty

We apply the following robustness concepts for this setting: Strict robustness, bounded uncertainty and ellipsoidal uncertainty. The parameter Γ describing the bounded uncertainty concept is chosen from the set $\{1, 2, \ldots, 10\}$. The parameter Ω

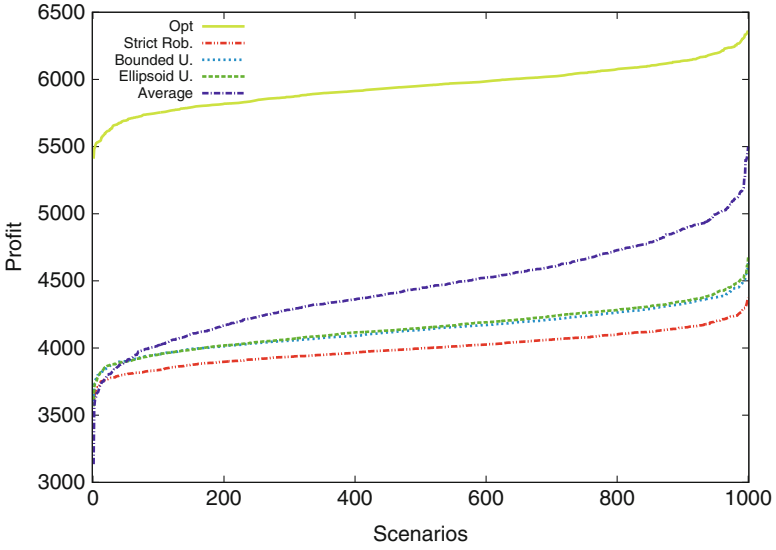


Fig. 7.6: The sampled scenario curve of an knapsack instance with 200 items and interval profit uncertainty. To generate the scenario curve, 1000 scenarios are sampled

defining the size of the ellipsoid used in the ellipsoidal uncertainty concept is chosen from the set $\{0.5, 1.0, 1.5, 2.0\}$. We present the price of robustness in Fig. 7.7 and the sampled scenario curve in Fig. 7.8.

We first consider Fig. 7.7. The strict robust solution is calculated under the assumption that every item has it highest possible weight. Hence, this solution is feasible for all possible parameter realizations, i.e. with 100 %. The average case solution, on the other hand, assumes that every item attains its average weight. Therefore, in almost 50 % of all parameter realizations the average case solution is infeasible, as the budget constraint is violated. Again, similar to the AC-WC curve the bounded and ellipsoidal uncertainty concepts generate interesting compromise solutions between worst and average case. The indicated curve is steep near the strict robust solution, which shows the fact that a small relaxation of the feasibility requirement can lead to a big improvement in the average case performance.

We now consider Fig. 7.8. For clarity we only show two solutions of the bounded and ellipsoidal uncertainty set, one for the smallest and one for the largest parameter used in the computation. The performance of an infeasible solution is set to 0. The first look at the average case solution reveals directly that the solution is feasible for about 50 % of the parameter realizations. Allowing infeasibility for few scenarios yields a significant improvement for all most all scenarios if the bounded or ellipsoidal uncertainty concept is used. If the parameter values of these concepts are chosen small, solutions are found that perform similar as the average case solution, but are feasible for more scenarios.

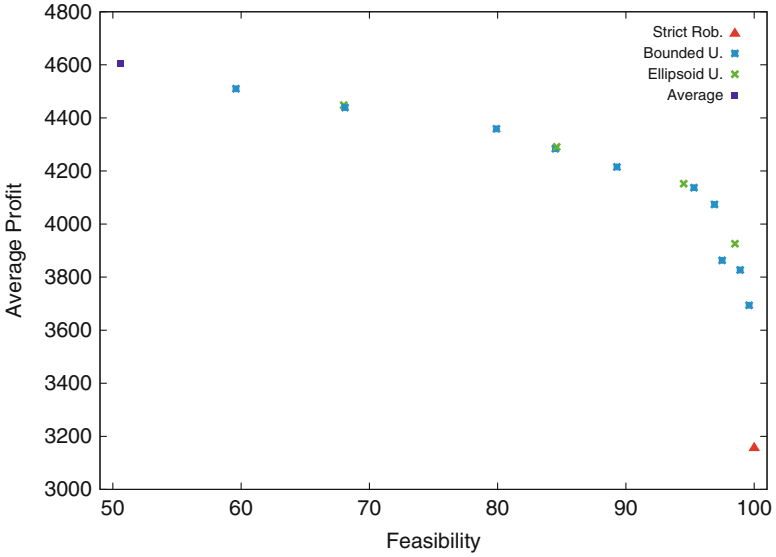


Fig. 7.7: Price of robustness for a knapsack problem with 200 items and interval uncertainty for profits and weights

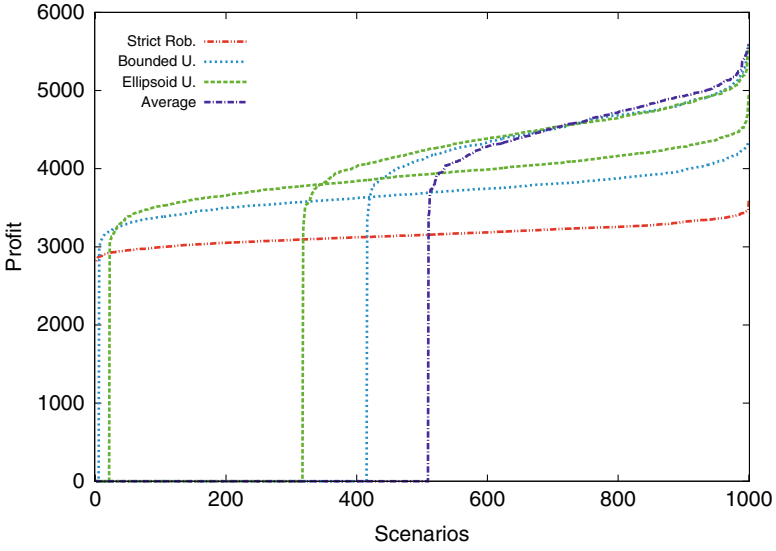


Fig. 7.8: The sampled scenario curve of a knapsack problem with 200 items and interval uncertainty for profits and weights. To generate the scenario curve, 1000 scenarios are sampled

Acknowledgements Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-13-1-3066. The U.S Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon.

References

1. Aissi, H., Bazgan, C., Vanderpooten, D.: Min–max and min–max regret versions of combinatorial optimization problems: a survey. Eur. J. Oper. Res. **197**(2), 427–438 (2009)
2. Averbakh, I.: Computing and minimizing the relative regret in combinatorial optimization with interval data. Discret. Optim. **2**(4), 273–287 (2005)
3. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. Math. Oper. Res. **23**(4), 769–805 (1998)
4. Ben-Tal, A., Nemirovski, A.: Robust solutions of linear programming problems contaminated with uncertain data. Math. Program. **88**, 411–424 (2000)
5. Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A.: Robust solutions of uncertain linear programs. Oper. Res. Lett. **25**, 1–13 (1999)
6. Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A.: Adjustable robust solutions of uncertain linear programs. Math. Program. A **99**, 351–376 (2003)
7. Ben-Tal, A., Ghaoui, L.E., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton/Oxford (2009)
8. Ben-Tal, A., Bertsimas, D., Brown, D.B.: A soft robust model for optimization under ambiguity. Oper. Res. **58**(4-Part-2), 1220–1234 (2010)
9. Bertsimas, D., Sim, M.: Robust discrete optimization and network flows. Math. Program. B **98**, 2003 (2003)
10. Bertsimas, D., Sim, M.: The price of robustness. Oper. Res. **52**(1), 35–53 (2004)
11. Bertsimas, D., Brown, D., Caramanis, C.: Theory and applications of robust optimization. SIAM Rev. **53**(3), 464–501 (2011)
12. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer Science & Business Media, New York (2011)
13. Buhmann, J.M., Mihalák, M., Srámek, R., Widmayer, P.: Robust optimization in the presence of uncertainty. In: Proceedings of the 4th conference on Innovations in Theoretical Computer Science, pp. 505–514. ACM (2013)
14. Carlsson, C., Fuller, R.: Fuzzy Reasoning in Decision Making and Optimization, vol. 82. Physica, Heidelberg (2012)
15. Chassein, A., Goerigk, M.: A bicriteria approach to robust optimization. Comput. Oper. Res. **66**, 181–189 (2015)
16. Fischetti, M., Monaci, M.: Light robustness. In: Ahuja, R.K., Möhring, R., Zaroliagis, C. (eds.) Robust and Online Large-Scale Optimization. Lecture Note on Computer Science, vol. 5868, pp. 61–84. Springer, Berlin (2009)
17. Goerigk, M., Schöbel, A.: Recovery-to-optimality: a new two-stage approach to robustness with an application to aperiodic timetabling. Comput. Oper. Res. **52**, Part A(0), 1–15 (2014)
18. Goerigk, M., Schöbel, A.: Algorithm engineering in robust optimization. In: Kliemann, L., Sanders, P. (eds.) Algorithm Engineering: Selected Results and Surveys. Lecture Notes in Computer Science, **9220** State of the Art. Springer (2016, to appear)
19. Hladık, M.: Interval linear programming: a survey. In: Mann, Z.A. (ed.) Linear Programming – New Frontiers in Theory and Applications, pp. 85–120. Nova Science Publishers, New York (2012)
20. Kalaï, R., Lamboray, C., Vanderpooten, D.: Lexicographic α-robustness: an alternative to min–max criteria. Eur. J. Oper. Res. **220**(3), 722–728 (2012)
21. Kellerer, H., Pferschy, U., Pisinger, D.: Knapsack Problems. Springer, Berlin (2004)
22. Kouvelis, P., Yu, G.: Robust Discrete Optimization and Its Applications. Kluwer Academic Publishers, Dordrecht (1997)

23. Liebchen, C., Lübbecke, M., Möhring, R., Stiller, S.: The concept of recoverable robustness, linear programming recovery, and railway applications. In: Robust and Online Large-Scale Optimization. Lecture Notes in Computer Science, vol. 5868, pp. 1–27. Springer, Berlin/Heidelberg (2009)
24. Monaci, M., Pferschy, U.: On the robust knapsack problem. SIAM J. Optim. **23**(4), 1956–1982 (2013)
25. Yaman, H., Karaşan, O.E., Pınar, M.Ç.: The robust spanning tree problem with interval data. Oper. Res. Lett. **29**(1), 31–40 (2001)

Chapter 8
Robust-Soft Solutions in Linear Optimization Problems with Fuzzy Parameters

Masahiro Inuiguchi

Abstract Linear optimization problems with fuzzy parameters were studied deeply and widely. Many of the approaches to fuzzy problems generate robust solutions. However, they were based on satisficing approaches so that the solutions do not maintain the optimality or suboptimality against the fluctuations in the coefficients. In this chapter, we describe a robust solution maintaining the suboptimality against the fluctuations in the coefficients. We formulate the problem as an extension of the minimax regret/maximin achievement rate problem and investigate a solution procedure based on a bisection method and a relaxation method. It is shown that the proposed solution procedure is created well so that both bisection and relaxation methods converge simultaneously.

8.1 Introduction

Due to the limit of available information, decision making problems often involve uncertainties. Traditionally two kinds of decision making problems under uncertainty have been studied: decision making problems under strict uncertainty and decision making problems under risk (see for example, [5]). In the former problems, the uncertainty is modelled by a set of possible situations where we do not know which situation is more probable than the others. In the latter problem, the uncertainty is modelled by a probability distribution. As an optimization technique under uncertainty, stochastic programming [18–20] were investigated. It treats decision

M. Inuiguchi (✉)
Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka, Japan
e-mail: inuiguti@sys.es.osaka-u.ac.jp

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_8

making problems under risk. However, recently, robust optimization [2] treating a kind of decision making problems under strict uncertainty is proposed and getting popular. In robust optimization, a solution which maintains the feasibility or the suboptimality against the parameter fluctuation in the given range is computed. Because of this property, the solution is considered a safe decision.

As a non-traditional model of uncertainty, fuzzy set theory [4, 21] was proposed and introduced into various fields. By fuzzy set theory, we can treat the vague restriction and goals of the decision maker (DM) on constraints and objectives as well as ambiguous coefficients in optimization problems [10, 19]. As we may treat the plausibility degree of a state of nature by a fuzzy set, we can formulate intermediate problems between decision making problems under strict uncertainty and decision making problems under risk. Fuzzy mathematical programming problems [10, 19] have been formulated so as to find a solution balanced between DM's aspiration and the robustness. Those formulations are based on a satisficing approach. Namely, the solution satisfies the given constraints and goals with a certain level of parameter fluctuation and is one of the best solutions in the balance between the robustness of given constraints and the possibility of achieving goals. However, the robustness in the sense that its objective function value is kept close to the optimal value against parameter fluctuation is not always high.

In this chapter, we describe optimizing approaches to linear programming problems with fuzzy objective function coefficients. An optimizing approach implies the formulations and solution methods obtaining robust solutions in the sense that their objective function values are kept close to the optimal value against parameter fluctuation. We introduce mainly two robust optimization approaches under softness: minimax regret type and maximin achievement rate type.

This chapter is organized as follows. In next section, blind spots in fuzzy programming approaches are shown by simple numerical examples. Two optimal solution concepts are given. We describe the weakness of those optimal solution concepts. In Sect. 8.3, solution concepts based on optimization approaches are described. Robust-soft optimal solutions maintaining suboptimality against the fluctuation in coefficients are defined in two ways. Solution algorithm under given fuzzy goals is investigated in Sect. 8.4. An acceleration technique in solving the subproblem is described in Sect. 8.5. In Sect. 8.6, a simpler solution algorithm is shown when fuzzy goals are not specified. Finally, in Sect. 8.7, we give some concluding remarks.

8.2 Blind Spots in Fuzzy Programming Approaches

8.2.1 Linear Program with Fuzzy Objective Function Coefficients

We treat the following linear programming problems with fuzzy objective function coefficients:

$$\text{maximize } \mathbf{c}^\top \mathbf{x}, \quad \text{subject to } \mathbf{x} \in F, \tag{8.1}$$

where $F = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ is bounded. $\mathbf{A} = (a_{ij})$ is an $m \times n$ constant matrix and $\mathbf{b} = (b_1, \ldots, b_m)^\top$. $\mathbf{x} = (x_1, \ldots, x_n)^\top$ is a decision variable vector. On the other hand, objective coefficient vector $\mathbf{c} = (c_1, c_2, \ldots, c_n)^\top$ is not known precisely but imprecisely. Namely, $\mathbf{c}$ takes a value in a possible range expressed by a bounded fuzzy set Γ of $\mathbb{R}^n$ with a membership function,

$$\mu_\Gamma(\mathbf{r}) = \min_{j=1,2,\ldots,p} \mu_{\Gamma_j}(\mathbf{d}_j^\top \mathbf{r}). \tag{8.2}$$

$\mathbf{r} \in \mathbb{R}^n$, Γ_j is a fuzzy number, i.e., a normal ($\exists r_j,\ \mu_{\Gamma_j}(r_j) = 1$), convex ($\mu_{\Gamma_j}$ is quasi-concave) and bounded fuzzy set on real line ($\lim_{r_j \to \infty} \mu_{\Gamma_j}(r_j) = \lim_{r_j \to -\infty} \mu_{\Gamma_j}(r_j) = 0$) with upper semi-continuous membership function μ_{Γ_j}. $\mathbf{d}_j \in \mathbb{R}^n$ is a constraint vector. The boundedness of Γ implies $p \geq n$, in other words, $\text{rank}\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_p\} = n$. Membership grade $\mu_\Gamma(\mathbf{r})$ can be understood as the possibility degree that $\mathbf{c} = \mathbf{r}$.

We define h-level sets and strong h-level sets by

$$[\Gamma]_h = \{\mathbf{r} \in \mathbb{R}^n \mid \mu_\Gamma(\mathbf{r}) \geq h\},\ [\Gamma_j]_h = \{\mathbf{r} \in \mathbb{R} \mid \mu_{\Gamma_j}(r) \geq h\},\ j = 1, 2, \ldots, p, \tag{8.3}$$

$$(\Gamma)_h = \{\mathbf{r} \in \mathbb{R}^n \mid \mu_\Gamma(\mathbf{r}) > h\},\ (\Gamma_j)_h = \{\mathbf{r} \in \mathbb{R} \mid \mu_{\Gamma_j}(r) > h\},\ j = 1, 2, \ldots, p. \tag{8.4}$$

We have

$$\begin{aligned} [\Gamma]_h &= \left\{\mathbf{r} \in \mathbb{R}^n \mid \mathbf{d}_j^\top \mathbf{r} \in [\Gamma_j]_h,\ j = 1, 2, \ldots, p\right\} \\ &= \left\{\mathbf{r} \in \mathbb{R}^n \mid \inf[\Gamma_j]_h \leq \mathbf{d}_j^\top \mathbf{r} \leq \sup[\Gamma_j]_h,\ j = 1, 2, \ldots, p\right\}, \end{aligned} \tag{8.5}$$

$$\begin{aligned} \text{cl}(\Gamma)_h &= \left\{\mathbf{r} \in \mathbb{R}^n \mid \mathbf{d}_j^\top \mathbf{r} \in \text{cl}(\Gamma_j)_h,\ j = 1, 2, \ldots, p\right\} \\ &= \left\{\mathbf{r} \in \mathbb{R}^n \mid \inf(\Gamma_j)_h \leq \mathbf{d}_j^\top \mathbf{r} \leq \sup(\Gamma_j)_h,\ j = 1, 2, \ldots, p\right\}, \end{aligned} \tag{8.6}$$

where cl$X$ is the closure of a set $X \subseteq \mathbb{R}^n$. An h-level set $[\Gamma]_h$ is depicted in Fig. 8.1.

Given a solution $\mathbf{x} \neq \mathbf{0}$, by the extension principle, its objective function value is given as a fuzzy set $Y(\mathbf{x})$ having the following membership function $\mu_{Y(\mathbf{x})}$:

$$\mu_{Y(\mathbf{x})}(y) = \sup_{\mathbf{c}} \{\mu_\Gamma(\mathbf{c}) : \mathbf{c}^\top \mathbf{x} = y\}. \tag{8.7}$$

Note that we have $Y(\mathbf{0}) = \{0\}$.

8.2.2 Solution Comparison by Objective Function Values

To treat linear programming problems with fuzzy coefficients, necessity measure N and possibility measure Π of a fuzzy set S are defined by

$$N_Q(S) = \inf_r \max(1 - \mu_Q(r), \mu_S(r)),\ \ \Pi_Q(S) = \sup_r \min(\mu_Q(r), \mu_S(r)), \tag{8.8}$$

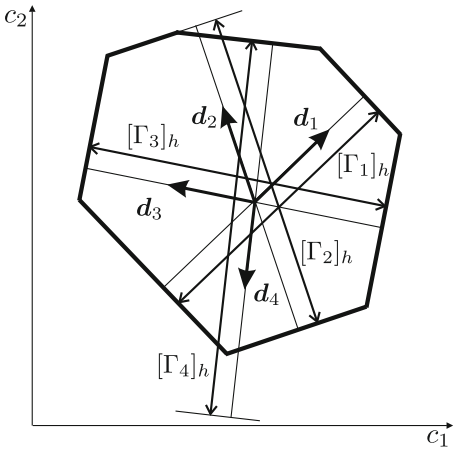


Fig. 8.1: h-Level set $[\Gamma]_h$

where μ_S and μ_Q are membership function of S and Q, respectively. $N_Q(S)$ and $\Pi_Q(S)$ evaluate to what extent the vague event expressed by fuzzy set S is necessary (certain) and possible under the possible range expressed by fuzzy set Q, respectively.

There are various ways to compare two fuzzy numbers Z_1 and $Z_2 \subseteq \mathbb{R}$. The following two indices are often used in the literature:

$$\mathrm{POS}(Z_1 \geq Z_2) = \sup_{r_1, r_2} \{\min(\mu_{Z_1}(r_1), \mu_{Z_2}(r_2)) : r_1 \geq r_2\}, \tag{8.9}$$

$$\mathrm{NES}(Z_1 \geq Z_2) = 1 - \sup_{r_1, r_2} \{\min(\mu_{Z_1}(r_1), \mu_{Z_2}(r_2)) : r_1 < r_2\}. \tag{8.10}$$

where μ_{Z_1} and μ_{Z_2} are membership functions of Z_1 and Z_2, and Z_1 and Z_2 are considered possible ranges of ambiguous numbers ζ_1 and ζ_2. Namely, we have a fuzzy set $Z_1 \times Z_2 \subseteq \mathbb{R}^2$ showing the possible ranges of (ζ_1, ζ_2) defined by a membership function $\mu_{Z_1 \times Z_2}(r_1, r_2) = \min(\mu_{Z_1}(r_1), \mu_{Z_2}(r_2))$. As an event we consider "ζ_1 is not smaller than ζ_2" which can be represented by a set (a binary relation) $\geq = \{(r_1, r_2) \in \mathbb{R}^2 \mid r_1 \geq r_2\}$. Then we have $\mathrm{POS}(Z_1 \geq Z_2) = \Pi_{Z_1 \times Z_2}(\geq)$ and $\mathrm{NES}(Z_1 \geq Z_2) = N_{Z_1 \times Z_2}(\geq)$. Namely, possibility degree $\mathrm{POS}(Z_1 \geq Z_2)$ shows to what extent Z_1 is possibly larger than or equal to Z_2. Similarly, Necessity degree $\mathrm{NES}(Z_1 \geq Z_2)$ shows to what extent Z_1 is necessarily larger than or equal to Z_2.

When Z_1 and Z_2 are closed intervals $[z_1^{\mathrm{L}}, z_1^{\mathrm{R}}]$ and $[z_2^{\mathrm{L}}, z_2^{\mathrm{R}}]$, respectively, we have

$$\mathrm{POS}(Z_1 \geq Z_2) = 1 \Leftrightarrow z_1^{\mathrm{R}} \geq z_2^{\mathrm{L}}, \qquad \mathrm{NES}(Z_1 \geq Z_2) = 0 \Leftrightarrow z_1^{\mathrm{L}} < z_2^{\mathrm{R}}. \tag{8.11}$$

Those equivalences remarkably show their meanings and difference.

A comparison index between two fuzzy numbers is often applied to the comparison of fuzzy objective function values discarding their interaction in literature. Next example demonstrates the inadequacy caused by the desertion of the interaction.

Example 1. Let $n = 2$ and $\Gamma = [1,2] \times [-2,-1]$. Namely, we consider a case when each objective function coefficient is given by a closed interval. Consider two feasible solutions $\mathbf{x}^1 = (2,1)^\top$ and $\mathbf{x}^2 = (3,1)^\top$. We have $Y(\mathbf{x}^1) = [0,3]$ and $Y(\mathbf{x}^2) = [1,5]$.

Let us apply the first equation of (8.11) discarding the interaction between $Z_1 = Y(\mathbf{x}^1)$ and $Z_2 = Y(\mathbf{x}^2)$. We obtain $\mathrm{POS}(Z_1 \geq Z_2) = 1$ which implies that the objective function value of $\mathbf{x}^1$ can be larger than or equal to that of $\mathbf{x}^2$. On the other hand, we have

$$\mathbf{c}^\top \mathbf{x}^1 = 2c_1 + c_2 < 3c_1 + c_2 = \mathbf{c}^\top \mathbf{x}^2, \ \ \forall c_1 \in [1,2], \ \forall c_2 \in [-2,-1]. \tag{8.12}$$

This insists that the objective function value of $\mathbf{x}^1$ can never be larger than or equal to that of $\mathbf{x}^2$. Because the realized values of c_1 and c_2 are common independent on the selection of a feasible solution of Problem (8.1), the latter result is correct. Therefore, the direct application of index $\mathrm{POS}(Z_1 \geq Z_2)$ is not adequate for the problem setting.

Similarly, from the second equation of (8.11), we obtain $\mathrm{NES}(Z_2 \geq Z_1) = 0$. This implies that there exists $(c_1,c_2)^\top \in \Gamma$ such that the objective function value of $\mathbf{x}^2$ is less than that of $\mathbf{x}^1$. However, this is neither true. As is shown in (8.12), for all $(c_1,c_2)^\top \in \Gamma$, the objective function value of $\mathbf{x}^2$ is larger than that of $\mathbf{x}^1$.

Now we emphasize the reason why indices defined by (8.9) and (8.10) do not work in the case of Example 1. Let ζ_1 and ζ_2 be variables taking values in Z_1 and Z_2, respectively. In the indices defined by (8.9) and (8.10), it is implicitly assumed that the possible range of ζ_2 does not depend on the realization of ζ_1 and also possible range of ζ_1 does not depend on the realization of ζ_2.

In Example 1, we set $Z_1 = Y(\mathbf{x}^1)$ and $Z_2 = Y(\mathbf{x}^2)$. Namely, they are possible ranges of $\zeta_1 = \mathbf{c}^\top \mathbf{x}^1$ and $\zeta_2 = \mathbf{c}^\top \mathbf{x}^2$, respectively. Both ζ_1 and ζ_2 depend on the realization of variable vector $\mathbf{c}$ taking a vector value in $\Gamma = [1,2] \times [-2,-1]$. Because of this fact, the implicit assumption in (8.9) and (8.10) does not hold. For example, when $\zeta_1 = \mathbf{c}^\top \mathbf{x}^1 = 0$, the possible realizations of $\mathbf{c} \in \Gamma$ are in

$$\{(c_1,c_2)^\top \in \mathbb{R}^2 : 2c_1 + c_2 = 0, 1 \leq c_1 \leq 2, -2 \leq c_2 \leq -1\} = \{(1,-2)\}. \tag{8.13}$$

Namely, from the information $\zeta_1 = 0$, in this case, we know the realization of $\mathbf{c}$ uniquely as $(1,-2)^\top$. It implies that ζ_2 is also uniquely known as $\zeta_2 = (1,-2)^\top \mathbf{x}^1 = 1$. Generally, when $\zeta_1 = q$, the possible range of ζ_2 is given by

$$\{3c_1 + c_2 : 2c_1 + c_2 = q, \ 1 \leq c_1 \leq 2, \ -2 \leq c_2 \leq -1\}. \tag{8.14}$$

This varies depending on ζ_1's realization q. Therefore, ζ_2 interacts with ζ_1. Similarly, ζ_1 interacts with ζ_2.

Since the implicit assumption of (8.9) and (8.10) does not hold, indices defined by (8.9) and (8.10) cannot be applied directly to the comparison between fuzzy objective function values. For the comparison between fuzzy objective function values, the following modified indices [7, 9] are adequate:

$$\mathrm{POS}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2) = \sup_{\mathbf{r}}\{\mu_\Gamma(\mathbf{r}) : \mathbf{r}^\top\mathbf{x}^1 \geq \mathbf{r}^\top\mathbf{x}^2\}, \tag{8.15}$$

$$\mathrm{NES}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2) = 1 - \sup_{\mathbf{r}}\{\mu_\Gamma(\mathbf{r}) : \mathbf{r}^\top\mathbf{x}^1 < \mathbf{r}^\top\mathbf{x}^2\}. \tag{8.16}$$

In literature, such desertion often appears when a fuzzy objective function is treated by a comparison of fuzzy numbers. Moreover, we note that under other interpretations of fuzzy coefficients, this discussion about the inadequacy is not valid. For example, when a fuzzy objective function is regarded as a collection of objective functions, e.g., a collection of utility functions of many decision makers, the above discussion cannot be applied.

Using $\mathrm{POS}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2)$ and $\mathrm{NES}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2)$ of (8.15) and (8.16), we can define fuzzy sets of necessary and possible non-inferior solutions by the following membership functions:

$$\mu_{NnS}(\mathbf{x}^2) = \begin{cases} 1 - \sup_{\mathbf{x}^1 \in F} \mathrm{POS}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2), & \text{if } \mathbf{x} \in F, \\ 0, & \text{if } \mathbf{x} \notin F, \end{cases} \tag{8.17}$$

$$\mu_{\Pi nS}(\mathbf{x}^2) = \begin{cases} 1 - \sup_{\mathbf{x}^1 \in F} \mathrm{NES}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2), & \text{if } \mathbf{x} \in F, \\ 0, & \text{if } \mathbf{x} \notin F. \end{cases} \tag{8.18}$$

We note that necessary non-inferior solution set NnS is defined by $\mathrm{POS}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2)$ while possible non-inferior solution set NnS is defined by $\mathrm{NES}(\mathbf{c}^\top\mathbf{x}^1 \geq \mathbf{c}^\top\mathbf{x}^2)$.

8.2.3 Necessity and Possibility Measure Optimization

When fuzzy goals G_1 and G_2 showing vaguely the required and desirable levels of objective function value are given, for example, Problem (8.1) can be treated as the following biobjective programming problem (see [10]):

$$\text{maximize } \left(N_{Y(\mathbf{x})}(G_1), \Pi_{Y(\mathbf{x})}(G_2)\right), \quad \text{subject to } \mathbf{x} \in F. \tag{8.19}$$

We assume that membership functions of G_1 and G_2 are non-decreasing.

When Γ is crisp and membership functions of G_1 and G_2 are increasing, Problem (8.19) is reduced to

$$\text{maximize } \left(\min_{\mathbf{c}\in\Gamma} \mathbf{c}^\top\mathbf{x}, \max_{\mathbf{c}\in\Gamma} \mathbf{c}^\top\mathbf{x}\right), \quad \text{subject to } \mathbf{x} \in F. \tag{8.20}$$

Completely optimal solutions to Problems (8.19) and (8.20), which maximize the both objective functions at the same time, have been regarded as the best solutions. However, as exemplified in the next example, such a complete optimal solution is not always the best solution.

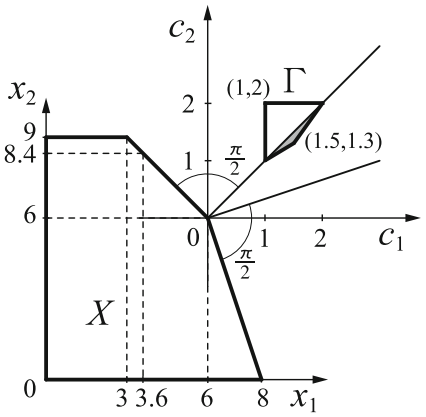


Fig. 8.2: The problem of Example 2

Example 2. Let us consider the following linear programming problem with uncertain objective coefficients:

$$\begin{aligned} &\text{maximize } c_1x_1 + c_2x_2, \\ &\text{subject to } x_1 + x_2 \le 12,\ 3x_1 + x_2 \le 24, \\ &\qquad x_2 \le 9,\ -x_1 \le 0,\ -x_2 \le 0, \end{aligned} \tag{8.21}$$

where $\mathbf{c} = (c_1, c_2)^\top$ is restricted by

$$\begin{aligned} \Gamma = \{(c_1, c_2,)^\top \mid &-4 \le 7c_1 - 5c_2 \le 4, \\ &2 \le -3c_1 + 5c_2 \le 9,\ 0 \le c_2 \le 2,\ 1 \le c_1 \le 3\}. \end{aligned} \tag{8.22}$$

For every $\mathbf{r} = (r_1, r_2) \in \Gamma$, we have $(1,1)^\top \le \mathbf{r} \le (2,2)^\top$, $(1,1)^\top \in \Gamma$ and $(2,2)^\top \in \Gamma$. The biobjective programming problem becomes

$$\text{maximize}(x_1 + x_2,\ 2x_1 + 2x_2),\ \text{ subject to } \mathbf{x} = (x_1, x_2)^\top \in F. \tag{8.23}$$

This problem has a completely optimal solution $\mathbf{x}^* = (6,6)^\top$.

The solution is illustrated in Fig. 8.2. The shaded area of Fig. 8.2 is the set of $\mathbf{c} \in \Gamma$ which makes $(6,6)^\top$ optimal. This shaded area is small relatively to Γ. From the viewpoint of optimality, $(6,6)^\top$ is not very robust because it easily fails to be optimal. However, this solution is robust in the sense that the objective function value is never less than $6 + 6 = 12$ as far as $\mathbf{c}$ fluctuates in $[1,2] \times [1,2]$.

8.3 Optimization Approaches

8.3.1 Possible and Necessary Optimal Solutions

Let $P(\mathbf{x}) = \{\mathbf{r} \mid \mathbf{r}^\top \mathbf{x} = \max_{\mathbf{y} \in F} \mathbf{r}^\top \mathbf{y}\}$. Possibly optimal solution set ΠS and necessarily optimal solution set NS are proposed for Problem (8.1) (see [11]). They are defined as fuzzy sets with the following membership functions, respectively:

$$\mu_{\Pi S}(\mathbf{x}) = \begin{cases} \sup\limits_{\mathbf{r} \in P(\mathbf{x})} \mu_\gamma(\mathbf{r}), & \text{if } \mathbf{x} \in F, \\ 0, & \text{if } \mathbf{x} \notin F, \end{cases} \tag{8.24}$$

$$\mu_{NS}(\mathbf{x}) = \begin{cases} \sup\limits_{\mathbf{r} \notin P(\mathbf{x})} 1 - \mu_\Gamma(\mathbf{r}), & \text{if } \mathbf{x} \in F, \\ 0, & \text{if } \mathbf{x} \notin F. \end{cases} \tag{8.25}$$

We have $\mu_{NS}(\mathbf{x}) > 0 \Rightarrow \mu_{\Pi S}(\mathbf{x}) = 1$. $\mu_{\Pi S}(\mathbf{x})$ and $\mu_{NS}(\mathbf{x})$ are called the possible optimality degree and necessary optimality degree of solution $\mathbf{x}$, respectively.

When a feasible solution is given, we will be interested in the degrees of $\mu_{\Pi S}$ and μ_{NS}. This topic is studied in [11] when $p = n$ and $\mathbf{d}_j = \mathbf{e}_j$, where $\mathbf{e}_j$ is a unit vector whose j-th component is one. The method is easily extended to the general case.

A solution such that $\mu_{NS}(\mathbf{x}) > 0$ (resp. $\mu_{\Pi S}(\mathbf{x}) > 0$) is called a necessarily (resp. possibly) optimal solution. In Example 2, the solutions on the line segment from $(6,6)^\top$ to $(3,9)^\top$ are possibly optimal solutions and there is no necessarily optimal solution. Generally, a necessarily optimal solution does not always exist but usually there are a lot of possibly optimal solutions. If a necessarily optimal solution exists, the solution is the most rational solution. Since F and Γ are bounded, any possibly optimal solution can be expressed as a convex combination of possibly optimal basic solutions. The number of possibly optimal solutions are finite because of the boundedness of F. An enumeration method of all possible optimal basic solutions together with possible optimality degrees $\mu_{\Pi S}$ is proposed in [8].

We note that the possibly and necessary optimal solutions sets equal to possibly and necessary non-inferior solution sets, i.e., we have $\Pi S = \Pi nS$ and $NS = NnS$.

8.3.2 Robust-Soft Optimal Solutions

Since in many cases no solution with positive necessary optimality degree exists, let us weaken the concept of the necessary optimality. To this end, we introduce the concept of soft optimality. If the objective function value of a feasible solution is slightly smaller than the optimal value, the solution can be regarded as a suboptimal solution. From this point of view, we define a suboptimal solution set $T(\mathbf{c})$ to a linear programming problem with an objective function vector $\mathbf{c}$ as a fuzzy set with a membership function,

$$\mu_{T(\mathbf{c})}(\mathbf{x}) = \begin{cases} \mu_{Dif}\left(\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{y} - \mathbf{c}^\top \mathbf{x}\right), & \text{if } \mathbf{x} \in F, \\ 0, & \text{if } \mathbf{x} \notin F, \end{cases} \tag{8.26}$$

or

$$\mu_{T(\mathbf{c})}(\mathbf{x}) = \begin{cases} \mu_{Rat}\left(\dfrac{\mathbf{c}^\top \mathbf{x}}{\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{y}}\right), & \text{if } \mathbf{x} \in F, \\ 0, & \text{if } \mathbf{x} \notin F, \end{cases} \tag{8.27}$$

where μ_{Dif} is assumed to be upper semi-continuous and non-increasing. μ_{Rat} is assumed to be upper semi-continuous, and non-decreasing if $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{x} > 0$ and non-increasing if $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{x} < 0$. While (8.26) is useful whenever the decision maker takes care of the difference from the optimal value, (8.27) is useful when $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{x} \neq 0$ and the decision maker takes care of the achievement rate based on the optimal value.

Using T, the necessarily soft optimal solution set NT is defined by

$$\mu_{NT}(\mathbf{x}) = \inf_{\mathbf{c}} \max\left(1 - \mu_{\Gamma}(\mathbf{c}), \mu_{T(\mathbf{c})}(\mathbf{x})\right). \tag{8.28}$$

It should be noted that NT with $T(\mathbf{c})$ defined by (8.27) is useful only when there exists $\mathbf{y} \in F$ such that $\min_{\mathbf{c}\in \mathrm{cl}(\Gamma)_h} \mathbf{c}^\top \mathbf{y} > 0$ for all $h \in [0,1)$ or when $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{y} < 0$ for all $\mathbf{c} \in (\Gamma)_0$. When we define a fuzzy set $V(\mathbf{x})$ of $\mathbf{c} \in \mathbb{R}^n$ for $\mathbf{x} \in \mathbb{R}^n$ by $\mu_{V(\mathbf{x})}(\mathbf{c}) = \mu_{T(\mathbf{c})}(\mathbf{x})$, we obtain $\mu_{NT}(\mathbf{x}) = N_{\Gamma}(V(\mathbf{x}))$, i.e., the necessarily soft optimal solution set NT is defined by using a necessity measure.

When $\mu_{Dif}(r)$ takes 1 for $r \leq 0$ and 0 for $r > 0$, fuzzy set NT is reduced to NS. Similarly, in the case of $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{x} > 0$, when $\mu_{Ra}(r)$ takes 1 for $r \geq 1$ and 0 for $r < 1$, fuzzy set NT is reduced to NS. In the case of $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{x} < 0$, when $\mu_{Ra}(r)$ takes 1 for $r \leq 1$ and 0 for $r > 1$, fuzzy set NT is reduced to NS.

8.4 Solution Algorithms Under Given Fuzzy Goals

When μ_{Dif} or μ_{Rat} is given by the decision maker, the best solution among NT is a solution with highest necessary soft optimality degree $\mu_{NT}(\mathbf{x})$. This problem is formulated as

$$\text{maximize} \ \ \mu_{NT}(\mathbf{x}). \tag{8.29}$$

This formulation was already proposed in [13]. While the objective function values of a solution $\mathbf{x}$ are independent of F in Problem (8.19), the objective function value depends on F in Problem (8.29). In this sense, both objective function values of Problem (8.19) are *independent* from other feasible solutions but the objective function value of Problem (8.29) is *depends* on others. Let $\hat{\mathbf{x}}$ and $\hat{h}$ be an optimal solution and the optimal value of Problem (8.29). Then we have

$$\forall \mathbf{c} \in (\Gamma)_{1-\hat{h}}, \ \mu_{T(\mathbf{c})}(\hat{\mathbf{x}}) \geq \hat{h}. \tag{8.30}$$

This implies that the suboptimality degree is guaranteed as at least $\hat{h}$ as far as $\mathbf{c}$ takes a value in $(\Gamma)_{1-\hat{h}}$. In this sense, Problem (8.29) produces a solution which is robust in suboptimality.

When $T(\mathbf{c})$ is defined by (8.26) and $\mathbf{d}_j = \mathbf{e}_j$, $j = 1,2,\ldots,p = n$, the equivalent problem and a solution algorithm based on bisection and relaxation methods are shown in [13]. In this chapter, we describe the result when $\mathbf{d}_j$, $j = 1,2,\ldots,p$ are general.

For the sake of simplicity, we consider the following three cases:

Case (I): $T(\mathbf{c})$ is defined by (8.26),

Case (II): $T(\mathbf{c})$ is defined by (8.27) under the assumption that there exists $\mathbf{y} \in F$ such that $\min_{\mathbf{c}\in(\Gamma)_h} \mathbf{c}^\top \mathbf{y} > 0$ for all $h \in (0,1]$,

Case (III): $T(\mathbf{c})$ is defined by (8.27) under the assumption that $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{y} < 0$ for all $\mathbf{c} \in (\Gamma)_0$

The procedure is the same among those three cases but subproblems are different. We note that in cases (II) and (III) we implicitly assume that $(\Gamma)_0$ is bounded.

We investigate (8.29) when $T(\mathbf{c})$ is defined by (8.26) and (8.27). Introducing an auxiliary variable h, from the upper semi-continuity of μ_{Rat}, Problem (8.29) is reduced to

$$\begin{cases} \text{maximize } h, \ \text{subject to } \mathbf{x} \in F, \ \mu_{Dif}\left(\max\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F}} (\mathbf{c}^\top \mathbf{y} - \mathbf{c}^\top \mathbf{x})\right) \geq h, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in case (I)}, \\ \text{maximize } h, \ \text{subject to } \mathbf{x} \in F, \ \mu_{Rat}\left(\min\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F,\ \mathbf{c}^\top \mathbf{y}>0}} \dfrac{\mathbf{c}^\top \mathbf{x}}{\mathbf{c}^\top \mathbf{y}}\right) \geq h, \quad \text{in case (II)}, \\ \text{maximize } h, \ \text{subject to } \mathbf{x} \in F, \ \mu_{Rat}\left(\max\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F}} \dfrac{\mathbf{c}^\top \mathbf{x}}{\mathbf{c}^\top \mathbf{y}}\right) \geq h, \quad \text{in case (III)}. \end{cases} \tag{8.31}$$

As h increases, $\mathrm{cl}(\Gamma)_{1-h}$ enlarges and thus, (1) the maximum value of $(\mathbf{c}^\top \mathbf{y} - \mathbf{c}^\top \mathbf{x})$ under $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$ and $\mathbf{y} \in F$ increases, (2) the minimum value of $\mathbf{c}^\top \mathbf{x}/\mathbf{c}^\top \mathbf{y}$ under $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$, $\mathbf{y} \in F$ and $\mathbf{c}^\top \mathbf{y} > 0$ decreases and (3) the maximum value of $\mathbf{c}^\top \mathbf{x}/\mathbf{c}^\top \mathbf{y}$ under $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$ and $\mathbf{y} \in F$ increases. Therefore, Problem (8.31) can be solved by a bisection method with respect to h with checking

$$\begin{cases} \mu_{Dif}\left(\min\limits_{\mathbf{x}\in F}\max\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F}}(\mathbf{c}^\top\mathbf{y}-\mathbf{c}^\top\mathbf{x})\right)\geq h, \text{ in case (I)},\\ \mu_{Rat}\left(\max\limits_{\mathbf{x}\in F}\min\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F,\ \mathbf{c}^\top\mathbf{y}>0}}\dfrac{\mathbf{c}^\top\mathbf{x}}{\mathbf{c}^\top\mathbf{y}}\right)\geq h, \text{ in case (II)},\\ \mu_{Rat}\left(\min\limits_{\mathbf{x}\in F}\max\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F}}\dfrac{\mathbf{c}^\top\mathbf{x}}{\mathbf{c}^\top\mathbf{y}}\right)\geq h, \text{ in case (III)}, \end{cases} \tag{8.32}$$

If (8.32) is satisfied, we know that the optimal value of Problem (8.31) is not less than h, and examine (8.32) again with an increased h. Otherwise, we know that the optimal value of Problem (8.31) is less than h and examine (8.32) again with a decreased h. Repeating this procedure, the possible range of the optimal value of Problem (8.31) reduced and we stop the procedure when the range becomes small enough.

In order to check the validity of (8.32), we should solve the min-max or max-min problem included in (8.32). Let us look into a solution method for these min-max and max-min problems.

Let $\mathbf{c}^j : (0,1] \to (\Gamma)_0$, $j = 1,2,\ldots,k$ be vector functions such that $\mathbf{c}^j(h) \in \mathrm{cl}(\Gamma)_h$, for all $h \in [0,1)$. These vector functions are generated through the algorithm proposed later in this chapter. Each of these function values is usually obtained at a vertex of $\mathrm{cl}(\Gamma)_h$ by solving a linear programming problem defined by an index set $Q_j = \{q_{j1}, q_{j2}, \ldots, q_{jn}\} \subseteq P = \{1,2,\ldots,p\}$ and a 0-1 vector $B_j = (\beta_{j1}, \beta_{j2}, \ldots, \beta_{jn})^\top \subseteq \{0,1\}^n$. We assume that $q_{j1} < q_{j2} < \cdots < q_{jn}$. Namely, given $h \in (0,1]$, the function value is obtained as $\mathbf{c}$-value of an optimal solution $(\mathbf{c}^\top, \delta_1, \delta_2, \ldots, \delta_n)^\top$ to the following linear programming problem:

$$\begin{aligned} &\text{minimize } \sum_{i=1}^{n}\delta_i,\\ &\text{subject to } \inf(\Gamma_{q_{ji}})_h + \delta_i = \mathbf{d}_{q_{ji}}^\top\mathbf{c} \leq \sup(\Gamma_{q_{ji}})_h, \text{ for } i \in N \text{ such that } \beta_{ji} = 0,\\ &\qquad \inf(\Gamma_{q_{ji}})_h \leq \mathbf{d}_{q_{ji}}^\top\mathbf{c} = \sup(\Gamma_{q_{ji}})_h - \delta_i, \text{ for } i \in N \text{ such that } \beta_{ji} = 1,\\ &\qquad \inf(\Gamma_q)_h \leq \mathbf{d}_q^\top\mathbf{c} \leq \sup(\Gamma_q)_h,\ q \in P\backslash Q_j\\ &\qquad \delta_i \geq 0,\ i = 1,2,\ldots,n, \end{aligned} \tag{8.33}$$

where $N = \{1,2,\ldots,n\}$.

When $\mathbf{y}^j \in F$, $j = 1,2,\ldots,k$ are given under fixed $h \in (0,1]$, a relaxation problem of the min-max/max-min problem in (8.32) is obtained as the following linear programming problem:

$$\begin{cases} \text{minimize } r, \text{ subject to } \mathbf{x} \in F, \ \mathbf{c}^j(1-h)^\top \mathbf{y}^j - \mathbf{c}^j(1-h)^\top \mathbf{x} \le r, \ j \in K, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in case (I)}, \\ \text{maximize } r, \text{ subject to } \mathbf{x} \in F, \ \dfrac{\mathbf{c}^j(1-h)^\top \mathbf{x}}{\mathbf{c}^j(1-h)^\top \mathbf{y}^j} \ge r, \ j \in K, \quad \text{in case (II)}, \\ \text{minimize } r, \text{ subject to } \mathbf{x} \in F, \ \dfrac{\mathbf{c}^j(1-h)^\top \mathbf{x}}{\mathbf{c}^j(1-h)^\top \mathbf{y}^j} \le r, \ j \in K, \quad \text{in case (III)}, \end{cases} \tag{8.34}$$

where $K = \{1,2,\ldots,k\}$ and the given solution $\mathbf{y}^j \in F$ satisfies $\mathbf{c}^j(1-h)^\top \mathbf{y}^j > 0$ when $T(\mathbf{c})$ is defined by (8.27) and $\max_{\mathbf{y}\in F} \mathbf{c}^\top \mathbf{y} > 0$ for all $\mathbf{c} \in (\Gamma)_0$. Note that the number of possible Q_j's is at most ${}_p\mathrm{C}_n = n!/(p!(n-p)!)$ and that the number of possible B_j's is 2^n. The value of $\mathbf{c}^j(h)$ is determined by solving Problem (8.33) for a given $h \in (0,1]$. Then the number of all possible vector functions $\mathbf{c}^j$ is $2^n n!/(p!(n-p)!)$. Because $\mathrm{cl}(\Gamma)_{1-h}$ is a polytope for each $h \in (0,1]$, any element $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$ can be represented by a convex combinations of the vertices of $\mathrm{cl}(\Gamma)_{1-h}$. Let $V(\mathrm{cl}(\Gamma)_{1-h})$ be the set of vertices of $\mathrm{cl}(\Gamma)_{1-h}$. Then we have $V(\mathrm{cl}(\Gamma)_{1-h}) = \{\mathbf{c}^j(1-h),\ j = 1,2,\ldots,2^n n!/(p!(n-p)!)\}$. Hence, when $k = 2^n n!/(p!(n-p)!)$, Problem (8.34) is equivalent to the min-max/max-min problem in (8.32).

Let $\mathbf{x}^0$ and r^0 be an optimal solution and the optimal value of Problem (8.34), respectively. Since Problem (8.34) is a relaxed problem, we should examine whether $\mathbf{x}^0$ is an optimal solution to the min-max/max-min problem in (8.32) or not under the fixed $h \in (0,1]$. This can be done by checking whether the optimal value of the following problem is not less than r^0:

$$\begin{cases} \text{maximize } \mathbf{c}^\top \mathbf{y} - \mathbf{c}^\top \mathbf{x}^0, \text{ subject to } \mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}, \ \mathbf{y} \in F, \text{ in case (I)}, \\ \text{minimize } \dfrac{\mathbf{c}^\top \mathbf{x}^0}{\mathbf{c}^\top \mathbf{y}}, \text{ subject to } \mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}, \ \mathbf{y} \in F, \ \mathbf{c}^\top \mathbf{y} > 0, \quad \text{in case (II)}, \\ \text{maximize } \dfrac{\mathbf{c}^\top \mathbf{x}^0}{\mathbf{c}^\top \mathbf{y}}, \text{ subject to } \mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}, \ \mathbf{y} \in F, \quad \text{in case (III)}, \end{cases} \tag{8.35}$$

where we note that constraint $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$ is represent by a system of linear inequalities because we have

$$\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h} \text{ if and only if } \inf(\Gamma_q)_{1-h} \le \mathbf{d}_q^\top \mathbf{c} \le \sup(\Gamma_q)_{1-h}, \ q \in P. \tag{8.36}$$

If the optimal value of Problem (8.35) is not greater/less than r^0, $\mathbf{x}^0$ is an optimal solution to the min-max/max-min problem in (8.32). Otherwise, we add $\mathbf{c}^{k+1} : (0,1] \to (\Gamma)_0$ which satisfies

$$\begin{cases} \mathbf{c}^{k+1}(1-h)^{\top}\mathbf{y}^j - \mathbf{c}^{k+1}(1-h)^{\top}\mathbf{x} > r^0, \ \exists j \in K, \text{ in case (I)}, \\ \dfrac{\mathbf{c}^{k+1}(1-h)^{\top}\mathbf{x}}{\mathbf{c}^{k+1}(1-h)^{\top}\mathbf{y}^j} < r^0, \ \exists j \in K, \text{ in case (II)}, \\ \dfrac{\mathbf{c}^{k+1}(1-h)^{\top}\mathbf{x}}{\mathbf{c}^{k+1}(1-h)^{\top}\mathbf{y}^j} > r^0, \ \exists j \in K, \text{ in case (III)}, \end{cases} \tag{8.37}$$

for the fixed $h \in (0,1]$ and update $k = k+1$. Such function $\mathbf{c}^{k+1}$ can be determined by using an optimal solution to Problem (8.35). Namely, there exists an optimal solution $\mathbf{c}$ to Problem (8.35) which has at least n independent $\mathbf{d}_q$ such that $\mathbf{d}_q^{\top}\mathbf{c} = \inf(\Gamma_q)_{1-h}$ or $\mathbf{d}_q^{\top}\mathbf{c} = \sup(\Gamma_q)_{1-h}$. This implies that we can find Q_{k+1} and B_{k+1} corresponding to $\mathbf{c}^{k+1}$ whose function value is obtained by solving (8.33) with substitution $j = k+1$.

From the above discussion, we obtain the following solution algorithm based on the bisection method and the relaxation procedure with an admissible error $\varepsilon > 0$.

Algorithm 1

Step 1. Select Q_1 and B_1 arbitrarily in order to define $\mathbf{c}^1 : (0,1] \to (\Gamma)_0$. Let $\mathbf{y}^1$ be an optimal solution to the following linear programming problem:

$$\text{maximize } \mathbf{c}^1(0.5)^{\top}\mathbf{y}, \quad \text{subject to } \mathbf{y} \in F. \tag{8.38}$$

Step 2. Set $h^{\mathrm{L}} = 0$, $h^{\mathrm{U}} = 1$, $k = 1$ and $\mathbf{x}^0 = \mathbf{y}^1$.

Step 3. Set $h = \frac{1}{2}(h^{\mathrm{L}} + h^{\mathrm{U}})$ and let $\mathbf{y}^{k+1}$ and r^k be an optimal solution and the optimal value of Problem (8.35), respectively.

Step 4. If $\mu_{Dif}(r^k) \geq h$ or $\mu_{Rat}(r^k) \geq h$ then update $h^{\mathrm{L}} = h$ and return to Step 3.

Step 5. If $h^{\mathrm{U}} - h^{\mathrm{L}} \leq \varepsilon$ then terminate the algorithm. If $h^{\mathrm{U}} \leq \varepsilon$ then there is no feasible solution $\mathbf{x}$ such that $\mu_{NT}(\mathbf{x}) > \varepsilon$ and otherwise the optimal solution is obtained as $\mathbf{x}^0$.

Step 6. Construct $\mathbf{c}^{k+1}$ from a pair (Q_{k+1}, B_{k+1}) corresponding to an optimal solution of the latest problem solved at Step 3. If there is no $j \in \{1,2,\ldots,k\}$ such that $\mathbf{c}^j = \mathbf{c}^{k+1}$ and $\mathbf{c}^j(1-h)^{\top}\mathbf{y}^j = \mathbf{c}^{k+1}(1-h)^{\top}\mathbf{y}^{k+1}$ then update $k = k+1$.

Step 7. Set $h = \frac{1}{2}(h^{\mathrm{L}} + h^{\mathrm{U}})$. In cases (I) and (III), obtain an optimal solution $(\mathbf{x}^*, r^*)$ to Problem (8.34) and go to Step 8. In case (II), update $\mathbf{y}^j$ as an optimal solution to $\text{maximize}_{\mathbf{y} \in F} \ \mathbf{c}^j(1-h)^{\top}\mathbf{y}$.

Step 8. If $\mu_{Dif}(r^*) < h$ or $\mu_{Rat}(r^*) < h$, then set $h^{\mathrm{U}} = h$ and return to Step 7. Otherwise, set $\mathbf{x}^0 = \mathbf{x}^*$ and return to Step 3.

In this algorithm, we use μ_{Dif} in case (I), and μ_{Rat} in cases (II) and (III). In case (II), we update $\mathbf{y}^j$ at Step 7 so that we have $\mathbf{c}^j(1-h)^{\top}\mathbf{y}^j > 0$, $j = 1,2,\ldots,k$. The existence of such $\mathbf{y}^j \in F$ is guaranteed by the assumption described in case (II). Furthermore, we do not solve the max-min problem in (8.32) at each fixed h but solve it simultaneously with optimizing h so as to obtain a solution of Problem (8.31).

To prove the convergence of Algorithm 1, we use the following proposition.

Proposition 1. *If h^{L} is not updated at Step 4 in iteration $k \geq 2$, for the pair $(\mathbf{c}^{k+1}, \mathbf{y}^{k+1})$ obtained by solving Problem (8.35) at Step 3, there is no $l \leq k$ such that $\mathbf{c}^l = \mathbf{c}^{k+1}$ and $\mathbf{c}^j(1-h)^\top \mathbf{y}^j = \mathbf{c}^{k+1}(1-h)^\top \mathbf{y}^{k+1}$.*

Proof. Assume there exists $l \leq k$ such that $\mathbf{c}^l = \mathbf{c}^{k+1}$ and $\mathbf{c}^j(1-h)^\top \mathbf{y}^j = \mathbf{c}^{k+1}(1-h)^\top \mathbf{y}^{k+1}$ for the pair $(\mathbf{c}^{k+1}, \mathbf{y}^{k+1})$ corresponding to Problem (8.35) solved at Step 3 when h^{L} is not updated at Step 4 in iteration $k \geq 2$. Since h^{L} is not updated at Step 4, r^k satisfies $\mu_{Dif}(r^k) < h$ in case (I), and $\mu_{Rat}(r^k) < h$ in cases (II) and (III). On the other hand, because $k \geq 2$, it has returned to Step 3 from Step 8 and thus r^* of the optimal solution $(\mathbf{x}^*, r^*)$ obtained at the last visit of Step 7 satisfies $\mu_{Dif}(r^*) \geq h$ in case (I), and $\mu_{Rat}(r^*) \geq h$ in cases (II) and (III). Therefore, we have

$$\mu_{Dif}(r^k) < \mu_{Dif}(r^*) \text{ in case (I) and } \mu_{Rat}(r^k) < \mu_{Rat}(r^*) \text{ in cases (II) and (III).} \quad (*)$$

If we fix $\mathbf{y}$ at a feasible solution in Problem (8.35), the problem becomes a linear programming problem or a linear fractional programming problem with a decision variable vector $\mathbf{c}$. From the theories of linear and linear fractional programming [1, 3], the optimal solution $\mathbf{c}$ exists at an extreme point of $\mathrm{cl}(\Gamma)_{1-h}$. Together with this fact, the assumption of the existence of $l \leq k$ such that $\mathbf{c}^l = \mathbf{c}^{k+1}$ and $\mathbf{c}^j(1-h)^\top \mathbf{y}^j = \mathbf{c}^{k+1}(1-h)^\top \mathbf{y}^{k+1}$ implies

$$r^k = \mathbf{c}^{k+1}(1-h)^\top \mathbf{y}^{k+1} - \mathbf{c}^{k+1}(1-h)^\top \mathbf{x}^0 = \mathbf{c}^l(1-h)^\top \mathbf{y}^l - \mathbf{c}^l(1-h)^\top \mathbf{x}^0.$$

Because r^* is the optimal value of Problem (8.34), we have $r^k \leq r^*$ in cases (I) and (III), and $r^k \geq r^*$ in case (II). This implies

$$\mu_{Dif}(r^k) \geq \mu_{Dif}(r^*) \text{ in case (I) and } \mu_{Rat}(r^k) \geq \mu_{Rat}(r^*) \text{ in cases (II) and (III).}$$

This contradicts $(*)$. □

From Proposition 1, in every iteration, h^{L} is updated or a new pair $(\mathbf{c}^{k+1}, \mathbf{y}^{k+1})$ is added. The update of h^{L} reduces the difference $h^{\mathrm{U}} - h^{\mathrm{L}}$ to half or to less than half. The number of pairs $(\mathbf{c}^j, \mathbf{y}^j)$ is finite because the number of pairs (Q_j, B_j) is finite and F is bounded. Hence Algorithm 8.4 terminates in a finite number of iterations.

8.5 Solving the Subproblem

In Algorithm 1, all problems other than Problem (8.35) are linear programming problems and solved easily. However, Problem (8.35) is neither a linear programming problem nor a concave/convex programming problem but a convex maximization/concave minimization problem. To solve this problem, several approaches such as two-phase method, outer approximation method, cutting hyperplane method

and so on were proposed (see [12, 14, 15, 17]) when Γ is a crisp set. In two-phase method, all possibly optimal extreme points $\mathbf{z}^l \in F$, $l = 1,2,\ldots,q$ such that $\mu_{\Pi S}(\mathbf{z}^l) > 0$ are enumerated before the execution of Algorithm 1 and at Step 3 of Algorithm 1 we solve q linear programming problems/linear fractional programming problems (8.35) with fixing $\mathbf{y} = \mathbf{z}^l$, $l = 1,2,\ldots,q$. In case 1, we may apply a post optimality technique of linear programming for the change of objective coefficient vector when we solve q linear programming problems (8.35) with fixing $\mathbf{y} = \mathbf{z}^l$, $l = 1,2,\ldots,q$. Therefore, those q problems are solved sequentially without reinitialization of simplex tableau.

On the other hand, in cases (II) and (III), we cannot apply this technique directly. In case (III), because we have $\mathbf{c}^\top \mathbf{x}^0 < 0$ and $\mathbf{c}^\top \mathbf{y} < 0$ for any $\mathbf{y} \in F$, we obtain

$$\min_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h} \\ \mathbf{y}\in F}} \frac{\mathbf{c}^\top \mathbf{y}}{\mathbf{c}^\top \mathbf{x}^0} = \frac{1}{\max\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h} \\ \mathbf{y}\in F,}} \dfrac{\mathbf{c}^\top \mathbf{x}^0}{\mathbf{c}^\top \mathbf{y}}}. \tag{8.39}$$

Applying the linear fractional programming technique, the minimization problem in the left-hand side of (8.39) is reduced to

$$\text{minimize } \hat{\mathbf{c}}^\top \mathbf{y}, \text{ subject to } \hat{\mathbf{c}}^\top \mathbf{x}^0 = -1,\ \mathbf{y} \in F,\ \frac{\hat{\mathbf{c}}}{t} \in \mathrm{cl}(\Gamma)_{1-h}, t \geq 0. \tag{8.40}$$

To this reduced problem, we can apply the post optimality technique and thus q linear fractional programming problems are solved efficiently. We note an optimal solution to Problem (8.35) is obtained as $(\hat{\mathbf{c}}/t, \mathbf{y})$ from the obtained optimal solution $(\hat{\mathbf{c}}, \mathbf{y}, t)$ to Problem (8.40).

In case (II), we cannot obtain a similar result to (8.39). This is because there is no guarantee that we have $\mathbf{c}^\top \mathbf{x}^0 > 0$ for all $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$ at Step 3. However, because of the assumption described in case (II), we have a solution $\mathbf{y} \in F$ such that $\mathbf{c}^\top \mathbf{y} > 0$ for all $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$. In order to ensure $\mathbf{c}^\top \mathbf{x}^0 > 0$ for all $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$, we can add $\mathbf{c}^{k+1} \in \mathrm{cl}(\Gamma)_{1-h}$ such that ${\mathbf{c}^{k+1}}^\top \mathbf{x}^0 < 0$, iteratively at Step 6 with the replacement of Step 3 by the following step:

Step 3'. Set $h = \frac{1}{2}(h^{\mathrm{L}} + h^{\mathrm{U}})$ and solve the following linear programming problem:

$$\text{minimize } {\mathbf{x}^0}^\top \mathbf{c}, \text{ subject to } \mathbf{c} \in \mathrm{cl}(C)_{1-h}. \tag{8.41}$$

If the optimal value is negative, let $\bar{\mathbf{c}}$ be the obtained optimal solution and $\mathbf{y}^{k+1}$ an optimal solution to

$$\text{maximize } \bar{\mathbf{c}}^\top \mathbf{y}, \text{ subject to } \mathbf{y} \in F, \tag{8.42}$$

and then go to Step 6. Otherwise, let $\mathbf{y}^{k+1}$ and r^k be an optimal solution and the optimal value of Problem (8.35), respectively.

Because F is bounded, we obtain some $\mathbf{x}^0$ such that $\mathbf{c}^\top \mathbf{x}^0 > 0$ for all $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$ in a finite iterations. If $\mathbf{c}^\top \mathbf{x}^0 > 0$ for all $\mathbf{c} \in \mathrm{cl}(\Gamma)_{1-h}$ is ensured, we have

$$\max_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F, \mathbf{c}^\top\mathbf{y}>0}} \frac{\mathbf{c}^\top\mathbf{y}}{\mathbf{c}^\top\mathbf{x}^0} = \frac{1}{\min\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h}\\ \mathbf{y}\in F,\ \mathbf{c}^\top\mathbf{y}>0}} \dfrac{\mathbf{c}^\top\mathbf{x}^0}{\mathbf{c}^\top\mathbf{y}}}, \tag{8.43}$$

Applying a linear fractional programming technique, the maximization problem in the left-hand side problem of (8.43) is reduced to a bilinear programming problem,

$$\begin{aligned} &\text{maximize } \hat{\mathbf{c}}^\top\mathbf{y},\\ &\text{subject to } \hat{\mathbf{c}}^\top\mathbf{x}^0 = 1,\ \mathbf{y}\in F,\ \frac{\hat{\mathbf{c}}}{t}\in \mathrm{cl}(\Gamma)_{1-h}. \end{aligned} \tag{8.44}$$

Let $(\hat{\mathbf{c}},\mathbf{y},t)$ be an optimal solution to Problem (8.44). Then solution $(\hat{\mathbf{c}}/t,\mathbf{y})$ is an optimal solution to Problem (8.35). The post optimality technique of linear programming problem can be applied to the reduced problem (8.44).

The transformations of the fractional programming problems in cases (II) and (III) are useful when the outer approximation method is used for solving Problem (8.35). By the numerical experiments reported in [14], the outer approximation method solves Problem (8.35) efficiently. An outer approximation algorithm is shown as follows.

Algorithm 2

Step 1. Initialize $p=0$ and obtain a polytope Y_0 such that $F\subseteq Y_0$.
Step 2. Enumerate all elements of $\Pi B(Y_p)$.
Step 3. Calculate $f(\mathbf{y})$ for all $\mathbf{y}\in\Pi B(Y_p)$.

In case (I): let $\mathbf{y}^p$ be a solution which maximizes $f(\mathbf{y})$ subject to $\mathbf{y}\in\Pi B(Y_p)$. Moreover, let $\bar{\mathbf{c}}^p$ be a $\mathbf{c}\in\Gamma$ such that $f(\mathbf{y}^p)=\mathbf{c}^{\mathrm{T}}(\mathbf{y}^p-\mathbf{x}^0)$.
In case (II): In case (II) let $\mathbf{y}^p$ be a solution which minimizes $f(\mathbf{y})$ subject to $\mathbf{y}\in\Pi B(Y_p)$. Moreover, let $\bar{\mathbf{c}}^p$ be a $\mathbf{c}\in\Gamma$ such that $f(\mathbf{y}^p)=\mathbf{c}^{\mathrm{T}}\mathbf{y}^p/\mathbf{c}^\top\mathbf{x}^0$.
In case (III): let $\mathbf{y}^p$ be a solution which maximizes $f(\mathbf{y})$ subject to $\mathbf{y}\in\Pi B(Y_p)$. Moreover, let $\bar{\mathbf{c}}^p$ be a $\mathbf{c}\in\Gamma$ such that $f(\mathbf{y}^p)=\mathbf{c}^{\mathrm{T}}\mathbf{y}^p/\mathbf{c}^\top\mathbf{x}^0$.

Step 4. In cases (I) and (III) if $f(\mathbf{y}^p)\le r^0$, terminate the algorithm with setting $r^k=r^0$. In case (II) if $f(\mathbf{y}^p)\ge r^0$, terminate the algorithm with setting $r^k=r^0$.
Step 5. If $\mathbf{y}^p\in F$, terminate the algorithm with setting $\mathbf{c}^k=\bar{\mathbf{c}}^p$, $\mathbf{z}^k=\mathbf{y}^p$ and $r^k=f(\mathbf{y}^p)$.
Step 6. Solve a linear programming problem,

$$\max_{\mathbf{y}\in F} \bar{\mathbf{c}}^{p\mathrm{T}}\mathbf{y}, \tag{8.45}$$

and let $\mathbf{w}^p$ be an optimal solution. Let Z be a set defined by constraints whose corresponding slack variables are nonbasic at the optimal solution $\mathbf{w}^p$.
Step 7. Update $Y_{p+1}=Y_p\cap Z$ and $p=p+1$. Return to Step 2.

In the algorithm above, $\Pi B(Y_p)$ is the set of all possibly extreme points with positive possible optimality degrees of Problem (8.1) where F is replaced with Y_p. $f(\mathbf{y})$ is

defined by

$$f(\mathbf{y}) = \begin{cases} \max\limits_{\mathbf{c}\in\Gamma}(\mathbf{c}^\top\mathbf{y} - \mathbf{c}^\top\mathbf{x}^0), & \text{in case (I)}, \\ \min\limits_{\mathbf{c}\in\Gamma} \dfrac{\mathbf{c}^\top\mathbf{x}^0}{\mathbf{c}^\top\mathbf{y}}, & \text{in case (II)}, \\ \max\limits_{\mathbf{c}\in\Gamma} \dfrac{\mathbf{c}^\top\mathbf{x}^0}{\mathbf{c}^\top\mathbf{y}}, & \text{in case (III)}. \end{cases} \tag{8.46}$$

Extreme points of $\Pi B(Y_{p+1})$ can be obtained easily from extreme points of $\Pi B(Y_p)$ (see [14]).

In [14], the outer approximation method and other possible solution methods for Problem (8.35) are described in case (I). Moreover the results of numerical experiments in comparison among possible solutions methods for Problem (8.35) are explained in [14].

8.6 Solution Algorithms Under Unknown Goals

The determination of μ_{Dif} or μ_{Rat} can be difficult in some situations. Instead of giving μ_{Dif} or μ_{Rat}, the decision maker may tell to what extent the fluctuation of coefficients should be taken care of. In this situation, the decision maker specifies $h^0 \in (0, 1]$ so that we consider all $\mathbf{c} \in (\Gamma)_{1-h^0}$. Under such a situation, we consider

$$\begin{cases} \text{minimize } q_{Dif}, \text{ subject to } \mu_{NT}(\mathbf{x}) \geq h^0, & \text{in case (I)}, \\ \text{minimize } |q_{Rat} - 1|, \text{ subject to } \mu_{NT}(\mathbf{x}) \geq h^0, & \text{in cases (II) and (III)}, \end{cases} \tag{8.47}$$

where we define

$$\mu_{Dif}(r) = \begin{cases} 1, \text{ if } r \leq q_{Dif}, \\ 0, \text{ if } r > q_{Dif}, \end{cases} \qquad \mu_{Rat}(r) = \begin{cases} 1, \text{ if } r \leq |q_{Rat} - 1|, \\ 0, \text{ otherwise}. \end{cases} \tag{8.48}$$

Those problems produce a robust solution. Let $\hat{\mathbf{x}}$ and $\hat{q}$ be an optimal solution and the optimal value of Problem (8.47). Then we have

$$\begin{cases} \forall \mathbf{c} \in (\Gamma)_{1-h^0}, \ \forall \mathbf{y} \in F, \ \mathbf{c}^\top\mathbf{y} - \mathbf{c}^\top\hat{\mathbf{x}} \leq \hat{q}, & \text{in case (I)}, \\ \forall \mathbf{c} \in (\Gamma)_{1-h^0}, \ \forall \mathbf{y} \in F, \ \left| \dfrac{\mathbf{c}^\top\hat{\mathbf{x}}}{\mathbf{c}^\top\mathbf{y}} - 1 \right| \leq \hat{q}, & \text{in cases (II) and (III)}. \end{cases} \tag{8.49}$$

In this section, we show a simpler solution procedure to Problem (8.47). Problem (8.47) is reduced to

$$\begin{cases} \text{minimize} \quad \max\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h^0} \\ \mathbf{y}\in F}} (\mathbf{c}^\top\mathbf{y} - \mathbf{c}^\top\mathbf{x}), \text{ subject to } \mathbf{x} \in F, & \text{in case (I)}, \\ \text{maximize} \quad \min\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h^0} \\ \mathbf{y}\in F,\ \mathbf{c}^\top\mathbf{y}>0}} \dfrac{\mathbf{c}^\top\mathbf{x}}{\mathbf{c}^\top\mathbf{y}}, \text{ subject to } \mathbf{x} \in F, & \text{in case (II)}, \\ \text{minimize} \quad \max\limits_{\substack{\mathbf{c}\in \mathrm{cl}(\Gamma)_{1-h^0} \\ \mathbf{y}\in F,\ \mathbf{c}^\top\mathbf{y}>0}}, \text{ subject to } \mathbf{x} \in F, & \text{in case (III)}. \end{cases} \tag{8.50}$$

In each case, this problem is the same as the min-max or max-min problem appeared in the argument of the membership function in (8.32). Thus, the discussion in Section 4 is valid also for Problem (8.50). Because h^0 is fixed, Problem (8.50) is easier than Problem (8.31). Based on the relaxation procedure, we have the following algorithm.

Algorithm 3

Step 1. Select $\mathbf{c}^1 : (0,1] \to (\Gamma)_0$. Let $\mathbf{y}^1$ be an optimal solution to a linear programming problem,

$$\text{maximize } {\mathbf{c}^1}^\top \mathbf{y}, \quad \text{subject to } \mathbf{y} \in F. \tag{8.51}$$

Set $k = 1$, $\mathbf{x}^0 = \mathbf{y}^1$, and $r^0 = 0$ in case (I) and $r^0 = 1$ in cases (II) and (III).

Step 2. Let $(\mathbf{c}^{k+1}, \mathbf{y}^{k+1})$ and r^k be an optimal solution and the optimal value of Problem (8.35) with $h = h^0$, respectively.

Step 3. If $r^k \le r^0 + \varepsilon$ in cases (I) and (III) and if $r^k \ge r^0 - \varepsilon$ in case (II) then terminate the algorithm. The optimal solution is obtained as $\mathbf{x}^0$.

Step 4. Update $k = k+1$. Update $(\mathbf{x}^0, r^0)$ with an optimal solution to the following problem:

$$\begin{cases} \text{minimize } r, \text{ subject to } \mathbf{x} \in F, \ {\mathbf{c}^j}^\top \mathbf{y}^j - {\mathbf{c}^j}^\top \mathbf{x} \le r, \ j \in K, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in case (I)}, \\ \text{maximize } r, \text{ subject to } \mathbf{x} \in F, \ \dfrac{{\mathbf{c}^j}^\top \mathbf{x}}{{\mathbf{c}^j}^\top \mathbf{y}^j} \ge r, \ j \in K, \quad \text{in case (II)}, \\ \text{minimize } r, \text{ subject to } \mathbf{x} \in F, \ \dfrac{{\mathbf{c}^j}^\top \mathbf{x}}{{\mathbf{c}^j}^\top \mathbf{y}^j} \le r, \ j \in K, \quad \text{in case (III)}. \end{cases} \tag{8.52}$$

Return to Step 2.

Example 3. Let us apply the approach under unknown goal to the problem of Example 2 with $h^0 = 0.5$. We define $T(\mathbf{c})$ by (8.27). In the problem, we can confirm that $\min_{\mathbf{c} \in (\Gamma)_0} \mathbf{c}^\top \mathbf{y} > 0$, for any $\mathbf{y} \in F$. Namely, we consider case (II). Since Γ is crisp in this problem, there is no difference by the choice of $h^0 \in (0,1]$. Setting $\varepsilon = 0.00001$, we applied Algorithm 3. The computation process is shown in Table 8.1. The obtained solution is

$$(x_1, x_2)^\top = (3.6, 8.4)^\top, \tag{8.53}$$

and its location is shown in Fig. 8.2. As shown in Fig. 8.2, reflecting the shape of Γ, i.e., the fact that Γ has a small right lower part, the obtained solution is located near an extreme point $(x_1, x_2) = (3, 9)$ rather than $(x_1, x_2) = (6, 6)$.

In order to see the correspondences between $\mathbf{c}^i \in \mathrm{cl}(\Gamma)_h$ and pair (Q_i, B_i) which are used in Algorithm 1, we note that $\mathbf{c}^1 = (1.5, 1.3)^\top$ is a solution to Problem (8.33) with $Q_1 = \{1,2\}$ and $B_1 = (1,0)$ and $\mathbf{c}^2 = (1,2)$ is a solution to Problem (8.33) with $Q_2 = \{3,4\}$ and $B_2 = (1,0)$.

Table 8.1: Computation process

Step 1 We select $\mathbf{c}^1(h_0) = (1.5, 1.3)^\top$. We obtain $\mathbf{y}^1 = (6,6)^\top$. Set $k = 1$, $\mathbf{x}^0 = \mathbf{y}^1$ and $r^0 = 1$.
Step 2 Solve Problem (8.35) with $h = h^0$. We obtain $(\mathbf{c}^2, \mathbf{y}^2) = ((1,2)^\top, (3,9)^\top)$ and $r^1 = 0.857143$.
Step 3 $r^1 = 0.857143 < r^0 - \varepsilon = 1 - 0.00001$. Continue.
Step 4 We update $k = 2$. We obtain $\mathbf{x}^0 = (3.6, 8.4)^\top$ and $r^0 = 0.971429$. Return to Step 2.
Step 2 Solve Problem (8.35) with $h = h^0$. We obtain $(\mathbf{c}^3, \mathbf{y}^3) = ((1.5, 1.3)^\top, (6,6)^\top)$ and $r^1 = 0.971429$.
Step 3 $r^1 = 0.971429 \geq r^0 - \varepsilon = 0.971429 - 0.00001$. Terminate the algorithm. The optimal solution is $\mathbf{x}^0 = (3.6, 8.4)^\top$

8.7 Concluding Remarks

In this chapter, we described robust optimization approaches to linear programming problems with fuzzy parameters. We explained the necessary care of the interaction between objective function values of two solutions when they are compared. The insufficiency of satisficing approaches is exemplified by a simple example. Two optimal solution sets are defined and their properties are briefly described. To overcome the weak points of two optimality concepts under uncertainty, robust-soft optimality concept is introduced. The necessarily soft optimal solution set is defined. Two suboptimal solution sets are considered and then solution approaches to two necessarily soft optimal solution sets are investigated. Both cases use the same main solution procedure although the subproblems are different. The solution procedure is based on a bisection method and a relaxation method and combined successfully so that both methods converge simultaneously. Nevertheless, the solution procedure is generally much more difficult than that by the satisficing approach. When fuzzy goals are unknown, we do not need to use the bisection method and the solution procedure becomes simpler. However, the reduced problem is still a non-convex optimization problem. We hope that global optimization techniques [6, 16] as well as computer technologies would be developed so that problems in optimization approaches would be solved in a practically acceptable computation time.

References

1. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms, 3rd edn. Wiley, Hoboken (2006)
2. Ben-Tal, A., Ghaoui, L.E., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
3. Dantzig, G.B.: Linear Programming and Extensions. Princeton University Press, Princeton (1998)
4. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic, New York (1980)
5. French, S.: Decision Theory: An Introduction to the Mathematics of Rationality. Ellis Horwood Limited, Chichester (1986)

6. Horst, R., Tuy, H.: Global Optimization: Deterministic Approaches, 3rd revised and enlarged edition. Springer, Berlin (1995)
7. Inuiguchi, M.: Multiple objective linear programming with fuzzy coefficient. In: Figueira, J., Greco, S., Ehrgot, M. (eds.) Multiple Criteria Decision Analysis: State of the Art Surveys, pp. 723–760, Springer, New York (2004)
8. Inuiguchi, M.: Enumeration of all possibly optimal vertices with possible optimality degrees in linear programming problems with a possibilistic objective function. Fuzzy Optim. Decis. Making. **3**, 311–326 (2004)
9. Inuiguchi, M., Kume, Y.: Extensions of fuzzy relations considering interaction between possibility distributions and an application to fuzzy linear program (in Japanese). Trans. Inst. Syst. Control Inf. Eng. **3**(4), 93–102 (1992)
10. Inuiguchi, M., Ramík, J.: Possibilistic linear programming: a brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Sets Syst. **111**(1), 3–28 (2000)
11. Inuiguchi, M., Sakawa, M.: Possible and necessary optimality tests in possibilistic linear programming problems. Fuzzy Sets Syst. **67**(1), 29–46 (1994)
12. Inuiguchi, M., Sakawa, M.: Minimax regret solutions to linear programming problems with an interval objective function. Eur. J. Oper. Res. . **86**(3), 526–536 (1995)
13. Inuiguchi, M., Sakawa, M.: Robust optimization under softness in a fuzzy linear programming problem. Int. J. Approx. Reason. **18**, 21–34 (1998)
14. Inuiguchi, M., Tanino, T.: On computation methods for a minimax regret solution based on outer approximation and cutting hyperplanes. Int. J. Fuzzy Syst. **3**(4), 548–557 (2001)
15. Inuiguchi, M., Higashitani, H., Tanino, T.: On computation methods of the minimax regret solution for linear programming problems with uncertain objective function coefficients. In: Proceedings of 1999 IEEE International Conference on SMC, Tokyo, vol.3, pp. 979–984 (1999)
16. Locatelli, M., Schoen, F.: Global Optimization: Theory, Algorithms, and Applications. Siam, Philadelphia (2013)
17. Mausser, H.E., Laguna, M.: A new mixed integer formulation for the maximum regret problem. Int. Trans. Oper. Res. **5**, 389–403 (1998)
18. Shapiro, A., Dentcheva, D., Ruszczynski, A.: Lectures on Stochastic Programming: Modeling and Theory, 2nd edn. Siam, Philadelphia (2014)
19. Słowinski, R., Teghem J. (eds.): Stochastic Versus Fuzzy Approaches to Multiobjective Mathematical Programming Under Uncertainty. Kluwer, Dordrecht (1990)
20. Stancu-Minasian, I.M.: Stochastic Programming with Multiple Objective Functions. D. Reidel Publishing Company, Dordrecht (1984)
21. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)

Chapter 9
Robust Machine Scheduling Based on Group of Permutable Jobs

Christian Artigues, Jean-Charles Billaut, Azzedine Cheref, Nasser Mebarki, and Zakaria Yahouni

Abstract This chapter presents the "group of permutable jobs" structure to represent set of solutions to disjunctive scheduling problems. Traditionally, solutions to disjunctive scheduling problems are represented by assigning sequence of jobs to each machine. The group of permutable jobs structure assigns an ordered partition of jobs to each machine, i.e. a group sequence. The permutation of jobs inside a group must be all feasible with respect to the problem constraints. Such a structure provides more flexibility to the end user and, in particular, allows a better reaction to unexpected events. The chapter considers the robust scheduling framework where uncertainty is modeled via a discrete set of scenarios, each scenario specifying the problem parameters values. The chapter reviews the models and algorithms that have been proposed in the literature for evaluating a group sequence with respect to scheduling objectives for a fixed scenario as well as the recoverable robust optimization methods that have been proposed for generating robust group sequence

C. Artigues (✉) • A. Cheref
CNRS, LAAS, 7 avenue du Colonel Roche, 31400 Toulouse, France

Univ de Toulouse, LAAS, F-31400 Toulouse, France
e-mail: artigues@laas.fr; acheref@laas.fr

J.-C. Billaut
Laboratoire d'Informatique de l'Université de Tours, 64 avenue jean Portalis, 37200 Tours, France
e-mail: jean-charles.billaut@univ-tours.fr

N. Mebarki
LUNAM, Université de Nantes, IRCCyN Institut de Recherche en Communication et Cybernétique de Nantes, UMR CNRS 6597, Nantes, France
e-mail: nasser.mebarki@univ-nantes.fr

Z. Yahouni
LUNAM, IRCCyN/MELT, Manufacturing Engineering Laboratory of Tlemcen, Tlemcen University, Tlemcen, Algeria
e-mail: zakaria.yahouni@irccyn.ec-nantes.fr

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_9

against scenario sets. The methods based on group sequences are compared with standard robust scheduling approaches based on job sequences. Finally, methods for exploiting group sequences in an industrial context are discussed and an experiment for human decision making in a real manufacturing system based on groups is reported.

List of Notations

- $\mathscr{J}$ set of jobs $\{J_j\}_{j=1,\ldots,n}$
- $\mathscr{M}$ set of machines $\{M_k\}_{k=1,\ldots,m}$
- $p_{k,j}$ duration of job $J_j \in \mathscr{J}$ on machine $M_k \in \mathscr{M}$
- p_j duration of job $J_j \in \mathscr{J}$ (one machine case)
- $C_{k,j}$ completion time of job $J_j \in \mathscr{J}$ on machine $M_k \in \mathscr{M}$
- C_j completion time of job $J_j \in \mathscr{J}$ (one machine case or $C_j = \max_{M_k \in \mathscr{M}} C_{k,j}$).
- r_j release date of job $J_j \in \mathscr{J}$
- d_j due date of job $J_j \in \mathscr{J}$
- $L_j = C_j - d_j$ lateness of job $J_j \in \mathscr{J}$
- $C_{\max} = \max_{J_j \in \mathscr{J}} C_j$ makespan
- $L_{\max} = \max_{J_j \in \mathscr{J}} L_j$ maximum lateness
- σ_{ERD} sequence of jobs according to the earliest release date rule
- σ_{EDD} sequence of jobs according to the earliest due date rule
- $\mathscr{S}$ set of uncertainty scenarios
- r_j^s release date of job $J_j \in \mathscr{J}$ under scenario $s \in \mathscr{S}$
- d_j^s due date of job $J_j \in \mathscr{J}$ under scenario $s \in \mathscr{S}$
- p_j^s duration of job $J_j \in \mathscr{J}$ under scenario $s \in \mathscr{S}$ (one machine case)
- $p_{k,j}^s$ duration of job $J_j \in \mathscr{J}$ on machine $M_k \in \mathscr{M}$ in scenario $s \in \mathscr{S}$
- $r_{k,j}^s$ release date of job $J_j \in \mathscr{J}$ on machine $M_k \in \mathscr{M}$ in scenario $s \in \mathscr{S}$
- $d_{k,j}^s$ due date of job $J_j \in \mathscr{J}$ on machine $M_k \in \mathscr{M}$ in scenario $s \in \mathscr{S}$
- $\mathscr{C}^s$ set of feasible schedules for scenario $s \in \mathscr{S}$
- $C_{k,j}$ completion time of job $J_j \in \mathscr{J}$ on machine $M_k \in \mathscr{M}$
- π_j^k index of the machine that precedes machine M_k in job J_j routing
- Ω_j^k index of the machine that follows machine M_k in job J_j routing.
- $L_{\max}^s(C)$ maximum lateness in scenario $s \in \mathscr{S}$ of a schedule $C \in \mathscr{C}^s$
- σ_i^k index of the job at position i on machine $M_k \in \mathscr{M}$ in job sequence σ
- $\mathscr{C}^s(\sigma)$ set of feasible schedules compatible with job sequence σ in scenario s
- $EC^s(\sigma)$ earliest feasible schedule compatible with job sequence σ in scenario s
- $L_{\max}^s(\sigma)$ maximum lateness of $EC^s(\sigma)$ in scenario s
- Σ set of feasible job sequences (scenario independent)
- G group sequence (ordered partition of jobs on each machine)
- ν_k number of groups on machine $M_k \in \mathscr{M}$
- G_i^k i^{th} group on machine $M_k \in \mathscr{M}$
- $\mathscr{G}$ the set of (scenario-independent) feasible group sequences
- $\Sigma(G)$ set of job sequences compatible with group sequence G

- $\mathscr{C}^s(G)$ set of schedules compatible with group sequence G in scenario s
- $\mathscr{EC}^s(G)$ set of earliest schedules compatible with group sequence G in scenario s
- g_j^k position of the group containing $O_{k,j}$
- $\theta_{k,j}$ lower bound for the earliest starting time of $O_{k,j}$ in a group sequence (head)
- $\chi_{k,j}$ lower bound on the earliest completion time of $O_{k,j}$ in a group sequence
- $\theta'_{k,j}$ lower bound on the time between $C_{k,j}$ and $C_{\max}$ (tail)
- $\gamma(G)$ lower bound on the completion time of group G
- $\underline{\tau}_{k,j}$ largest earliest start time of $O_{k,j}$ in a group sequence
- $\underline{C}_{k,j}$ largest earliest completion time of $O_{k,j}$ in a group sequence
- $\overline{\tau}_{k,j}$ smallest latest start time of $O_{k,j}$ in a group sequence
- $\overline{C}_{k,j}$ smallest latest start time of $O_{k,j}$ in a group sequence
- $Prec_j^s$ set of jobs that precede J_j according to the ERD rule in scenario s
- $Succ_j^s$ set of jobs that succeed to J_j according to the ERD rule in scenario s
- $y_{j,q}$ binary variable equal to 1 if job J_j is in the group at position q in the group sequence
- $x_{j,q}$ binary variable equal to 1 if job J_j is at position q in the job sequence
- μ, α, β, ω parameters for instance generation
- $m_{\text{seq}}(O_{k,j})$ free sequential margin of an operation $O_{k,j}$ in a group sequence
- $m_{\text{sn}}(O_{k,j})$ net margin of an operation $O_{k,j}$ in a group sequence
- $m_{\text{sg}}(O_{k,j})$ group margin of an operation $O_{k,j}$ in a group sequence

9.1 Introduction to Scheduling and Robust Scheduling

In this section, disjunctive scheduling problems are defined and notations are introduced. The standard solution representations based on job sequences and disjunctive graphs are reviewed. Robustness definitions are given and the ways robustness can be tackled are also presented.

9.1.1 Scheduling Problems

A scheduling problem consists in defining a set of a start times for a set of tasks that share common resources, taking into account specific time constraints (such as deadlines), and with the aim to optimize an objective function. Sometimes, one also has to decide which resources will perform each task. A schedule is a solution to a scheduling problem (see [8] for a global overview of scheduling problems).

Scheduling problems can be found in all types of organizations or systems. The most famous application domain is the production industry, where scheduling problems take an important place for the production management. Other classical scheduling problems are encountered in computer systems, project management, timetabling, etc. More recent application domains appear, particularly for treating requests in big data environment [4, 5], in hospital environments [14], in rail companies [11], etc.

The theory of scheduling was developed together with the theory of complexity. The methods for solving scheduling problems come from the field of combinatorial optimization, with exact solution methods for finding one optimal solution to the problem and approximate solution methods for finding solutions that are as good as possible.

We denote by $\mathscr{J}$ the set of n jobs to schedule. A job is either a single operation, or is composed by several operations. In this case, a graph allows to define the precedence relations between the operations. In the case of a shop scheduling problem, this graph is called the route or the routing of the job, and it is generally a chain. A set $\mathscr{M}$ of m resources is available for performing the operations. Each resource (called machine in the case of a shop scheduling problem) is supposed to be ready at time 0, and can perform at most one operation at a time. To each operation is associated a performing resource in $\mathscr{M}$ and a processing time, denoted by $p_{k,j}$ for processing time of job J_j on machine M_k (index k is omitted when there is only one resource). It is assumed that the operations are performed without interruption, and that preemption is not allowed.

A schedule is completely characterized by the definition of a vector of starting times, or equivalently by a vector of completion times. We denote by $C_j = \max_{M_k \in \mathscr{M}} C_{k,j}$ the completion time of the last operation of job J_j, $C_{k,j}$ denoting the completion time of job J_j on machine M_k.

The quality of a schedule is given by a performance measure, based on the jobs completion times. The most common objective function is the makespan, denoted by $C_{\max}$ and defined by $C_{\max} = \max_{J_j \in \mathscr{J}} C_j$. If the job J_j is supposed to be finished at date d_j, called the due date of J_j, we define by L_j the lateness of J_j, with $L_j = C_j - d_j$, and $L_{\max} = \max_{J_j \in \mathscr{J}} L_j$, also called the maximum lateness, is another important performance measure for a schedule. Other performance measures can be defined for a schedule, please refer to [8] or [7] for a complete overview of scheduling problems.

In the rest of this chapter, we will consider only two scheduling problems: a scheduling problem with the environment composed by a single machine, and a scheduling problem with a shop environment, where the routes of the jobs are different (also called a *job shop* environment). These two problems are illustrated in the following example.

Example 1: Single Machine Environment

Let consider a single machine environment, and $n = 5$ jobs to schedule. To each job J_j is associated a processing time p_j, a due date d_j, and a release time r_j, which is a date before which a job cannot start. The objective is to schedule these jobs so that the maximum lateness $L_{\max}$ is as small as possible. The data set is given in Table 9.1.

Suppose that the jobs are scheduled in the non decreasing order of the release times. We obtain sequence σ_{ERD}. Suppose that the jobs are scheduled in the non decreasing order of the due dates (also called EDD rule for Earliest Due Date first), we obtain sequence σ_{EDD}. The sequences and their evaluations are given in Fig. 9.1.

Table 9.1: Instance for the single machine problem

J_j	J_1	J_2	J_3	J_4	J_5
r_j	0	7	3	4	3
p_j	3	4	1	2	4
d_j	3	14	4	6	10

Notice that this problem is NP-hard in the strong sense, i.e. no polynomial time algorithm can be proposed for solving the problem to optimality, unless $P = NP$ [13].

Example 2: Job Shop Environment

Let consider now a job shop environment, and $n = 3$ jobs to schedule on $m = 3$ machines and the data given in Table 9.2. To each job J_j are associated exactly 3 operations (one per machine). The objective is to schedule the jobs so that the makespan $C_{\max}$ is as small as possible.

Table 9.2: Instance for the job shop problem

	Machine (duration)		
J_1	$M_1(6)$	$M_2(3)$	$M_3(7)$
J_2	$M_3(8)$	$M_1(6)$	$M_2(4)$
J_3	$M_1(5)$	$M_3(5)$	$M_2(6)$

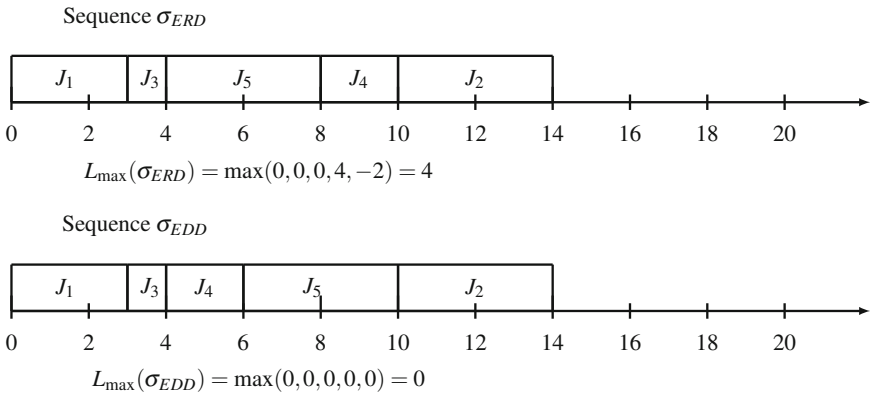


Fig. 9.1: Gantt representation of sequences σ_{ERD} and σ_{EDD} and their evaluations

Figure 9.2 represents a feasible solution of the problem with a makespan equal to 27. Note that the sequences of the jobs on each machine can be different from each other.

9.1.2 Robustness in Scheduling

Robustness considerations receive more and more attention in the literature [6] because in real life situations, unexpected events and uncertainty of the data are challenging the expected plans, making then unusable, sometimes very quickly. This is the reason why a lot of practitioners prefer a robust solution with a lower quality to a vulnerable solution with optimal quality.

Several definitions of the robustness can be found in the literature. We refer to the book by Kouvelis and Yu [15], where a chapter is devoted to robust scheduling problems. As in [15], we consider that there is a significant data uncertainty and the aim is to propose an algorithm returning a solution that hedges against the worst contingency that may arise. Such an approach is called a *robust approach*.

A scenario based approach is used to model the data uncertainty. Each scenario is a data set corresponding to a potential realization. Several scenarios are defined. More formally, let us denote by $\mathscr{S}$ the set of scenarios and s one scenario in $\mathscr{S}$.

We now informally illustrate the concept of robustness on the two scheduling problems that we consider.

Example 1: Single Machine Environment

Let us consider the single machine scheduling problem. We denote by r_j^s, p_j^s and d_j^s the release time, the processing time and the due date of J_j under scenario s. We assume that the previous data set is scenario number 1, and we add a new scenario.

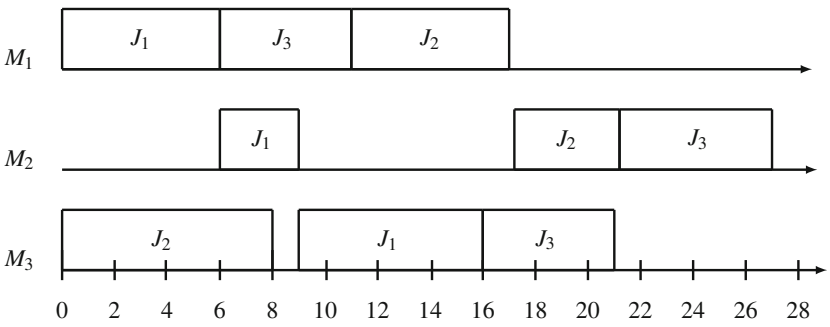


Fig. 9.2: Gantt representation of a solution to the job shop problem

The data set is now given in Table 9.3. Figure 9.1 presents the solutions for the first scenario, we present the sequences σ_{ERD} and σ_{EDD} for the second scenario in Fig. 9.3.

Table 9.3: Instance for the single machine problem with two scenarios

$s=1$	J_1	J_2	J_3	J_4	J_5	$s=2$	J_1	J_2	J_3	J_4	J_5
r_j^1	0	7	3	4	3	r_j^2	3	3	0	1	7
p_j^1	3	4	1	2	4	p_j^2	2	5	1	3	3
d_j^1	3	14	4	6	10	d_j^2	6	11	2	5	14

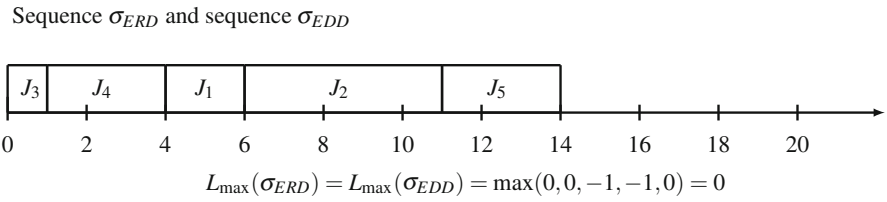


Fig. 9.3: Gantt representation of sequences σ_{ERD} and σ_{EDD} and their evaluations for the second scenario

We can see that for the second scenario, sequences σ_{ERD} and σ_{EDD} are the same: $(J_3, J_4, J_1, J_2, J_5)$. We can say that sequence σ_{EDD} is more robust than sequence σ_{ERD} because the worse evaluation of σ_{EDD} is 0 and the worse evaluation of σ_{ERD} is 4.

Example 2: Job Shop Environment

A second scenario is considered for the job shop problem. The two scenarios are presented in Table 9.4. Only the jobs durations are changed, not the routing of the jobs. Let $p_{k,j}^s$ denote the duration of job J_j on machine M_k in scenario s. If we keep the same sequence on each machine as before and schedule the operations as early as possible, we now obtain the solution represented in Fig. 9.4. The makespan is now equal to 29, which is the worst case makespan on the job sequences.

Table 9.4: Instance for the job shop problem to do

$s=1$	performing machine (duration)			$s=2$	performing machine (duration)		
J_1	$M_1(6)$	$M_2(3)$	$M_3(7)$	J_1	$M_1(8)$	$M_2(2)$	$M_3(5)$
J_2	$M_3(8)$	$M_1(6)$	$M_2(4)$	J_2	$M_3(10)$	$M_1(6)$	$M_2(5)$
J_3	$M_1(5)$	$M_3(5)$	$M_2(6)$	J_3	$M_1(6)$	$M_3(5)$	$M_2(4)$

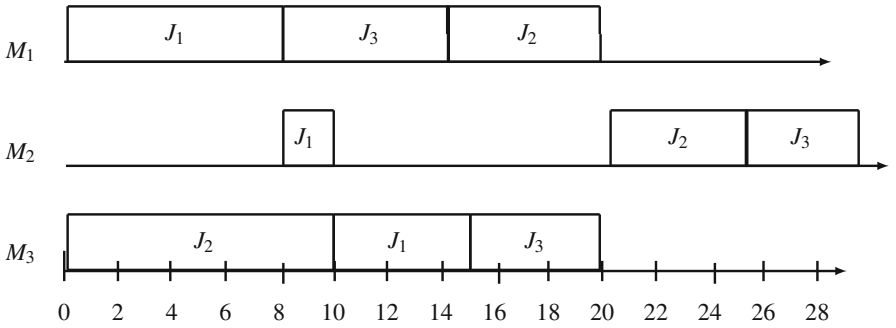


Fig. 9.4: Gantt representation of the same solution to the job shop problem with scenario $s = 2$

9.1.3 Feasible Schedules and the Absolute Robustness Problem

We consider in this section an integration of the job shop scheduling problem and of the single machine problem by defining operation release dates and due dates in the job shop model. More precisely, $r^s_{k,j} \geq 0$ is the release date of the operation of job J_j on machine M_k in scenario s and $d^s_{k,j}$ is the due date of the operation of job J_j on machine M_k in scenario s.

For a scenario $s \in \mathscr{S}$, a feasible solution is a feasible schedule, given by a completion time $C_{k,j}$ of each job J_j on each machine M_k that satisfies:

- Operations release date constraints:

$$C_{k,j} \geq r^s_{k,j} + p^s_{k,j} \quad \forall J_j \in \mathscr{J}, \forall M_k \in \mathscr{M} \tag{9.1}$$

- Jobs routing constraints:

$$C_{k,j} \geq C^s_{\pi^k_j,j} + p^s_{k,j} \quad \forall J_j \in \mathscr{J}, \forall M_k \in \mathscr{M}, \pi^k_j \neq 0 \tag{9.2}$$

  where π^k_j denotes the machine that precedes machine M_k in job J_j routing, with $\pi^k_j = 0$ indicating that machine M_k is the first machine in the routing of job J_j,
- Non-overlapping (also called disjunctive) constraints:

$$C_{k,j} \geq C_{k,i} + p^s_{k,j} \vee C_{k,i} \geq C_{k,j} + p^s_{k,i} \quad \forall J_i, J_j \in \mathscr{J}, i < j, \forall M_k \in \mathscr{M} \tag{9.3}$$

We denote by $\mathscr{C}^s$ the set of feasible schedules for scenario s, i.e. the set of vectors $(C_{k,j})_{M_k \in \mathscr{M}, J_j \in \mathscr{J}}$ that satisfy constraints (9.1)–(9.3). The maximum lateness of a schedule $C \in \mathscr{C}^s$ is given by

$$L^s_{\max}(C) = \max_{J_j \in \mathscr{J}} C_{k,j} - d^s_{k,j}. \tag{9.4}$$

The standard absolute robustness problem (AR) as defined in [15] can now be stated as

$$(AR) \min_{C \in \cap_{s \in \mathscr{S}} \mathscr{C}^s} \max_{s \in \mathscr{S}} L^s_{\max}(C) \tag{9.5}$$

9.1.4 The Standard Solution Representation for (Robust) Disjunctive Scheduling

We note that a schedule may be feasible for a scenario and infeasible for another one. In robust scheduling, it is convenient to consider the concept of job sequence that allows to represent compactly for each scenario a family of feasible schedules. The determination of the feasibility of a job sequence and the computation of the scenario-dependent schedules can be both supported by the classical disjunctive graph representation of the problem [22], defined as follows.

The disjunctive graph has the same vertices and arcs for all scenario but the weights of the arcs are scenario dependent. The disjunctive graph has $nm+2$ vertices, with a vertex jk per operation , for $j = 1,\ldots,n$ and for $k = 1,\ldots,m$ plus dummy start and end vertices 0 and $nm+1$. The disjunctive graph is a 2-graph that contains precedence arcs and disjunctive arcs. For a scenario s, we define a precedence arc $(0, jk)$ between vertex 0 and vertex jk valuated by $r^s_{k,j}$ for each release date constraint (9.1). We define a precedence arc between vertex $j\pi^k_j$ and jk, valuated by $p^s_{k,j}$ for each precedence constraint (9.2). For each disjunctive constraint (9.3), we define two opposite disjunctive arcs, one from node ik to node jk valuated by $p^s_{k,j}$ and one from node jk to node ik valuated by $p^s_{k,i}$. Last we have an arc from each node jk such that $\nexists k', \pi^{k'}_j = k$ and node $nm+1$ valuated by $p^s_{k,j} - d^s_{k,j}$.

A complete selection of the disjunctive graph is, for each pair of disjunctive arcs, the selection of a single arc (and the removal of the opposite one). A complete selection is acyclic if there is no cycle in the graph issued from the selection. Once an acyclic selection is obtained, it defines m total orders of the set of jobs via the selected disjunctive arcs. Hence an acyclic selection can be associated with a job sequence $\sigma = (\sigma^k_i)^{k=1,\ldots,m}_{i=1,\ldots,n}$ where σ^k_i gives the index of the job sequenced at position i on machine M_k. Notice that in the job shop problem a "job sequence" designates in fact a set of job sequences (one per machine). There is a one-to-one mapping between the acyclic selections and the feasible job sequences. Note that in the single machine problem, all job sequences are feasible, which yields to $n!$ job sequences. For the job shop problem, only a subset of job sequences are feasible due to precedence constraints, so we have at most $mn!$ feasible job sequences.

An acyclic complete selection/job sequence σ represents for each scenario a family of feasible schedules $\mathscr{C}^s(\sigma)$ given by the infinite set of potentials in the graph issued from the selection. In this family, a dominant schedule with respect to the $L_{\max}$ objective function is the earliest schedule $EC^s(\sigma)$ such that $EC^s_{k,j}(\sigma)$ is the length of the longest path from vertex 0 to vertex kj in the graph issued from the selection , the corresponding $L^s_{\max}(\sigma)$ is the longest path between 0 and $nm+1$.

Note that given a complete selection, the earliest schedule as well as the cycle determination can be computed in $O(nm)$ by topological sorting and Bellman-Ford algorithm.

Let Σ denote the set of feasible job sequences, which is independent of the scenarios. For a fixed scenario s the standard job shop scheduling problem can be defined as

$$\min_{\sigma \in \Sigma} L^s_{\max}(\sigma)$$

Similarly the absolute robust job shop scheduling problem can be defined as the search for a feasible job sequence that minimizes the worst case maximum lateness

$$(AR) \min_{\sigma \in \Sigma} \max_{s \in \mathscr{S}} L^s_{\max}(\sigma)$$

In terms of robust scheduling, the job sequence σ represents the first stage decision variables that can be taken in advance without knowledge of the realized scenario, while the completion times are the second-stage decision variables that are adjusted according to the realized scenario by picking a schedule in set $\mathscr{C}^s$. In case of an objective function defined as the sum or the maximum of non decreasing job individual functions of the completion time (also called a regular objective function), the earliest schedule $EC^s(\sigma)$ dominates all other schedules of $\mathscr{C}^s$.

Example 1: Single Machine Environment

In Fig. 9.5, we give the disjunctive graph representation of the one-machine problem instance. Because there is no routing constraint between the jobs, the disjunctive arcs form a clique. In Fig. 9.6 we give the conjunctive graph corresponding to the feasible job sequence given by σ_{ERD} for scenario $s = 2$.

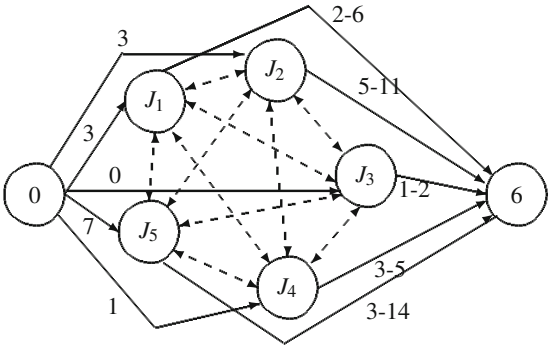


Fig. 9.5: Disjunctive graph of the single machine scheduling problem

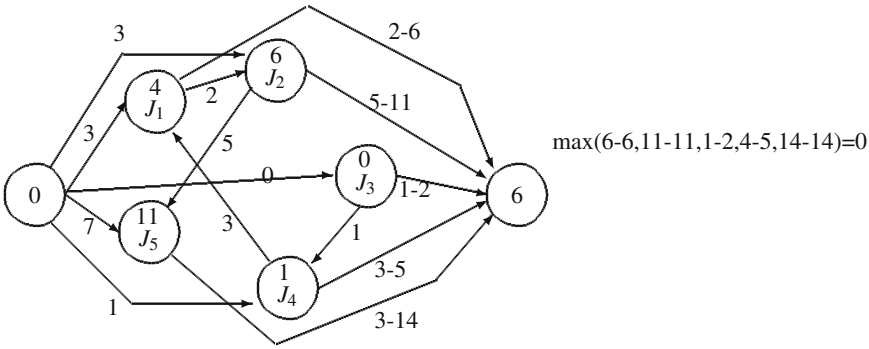


Fig. 9.6: Conjunctive graph corresponding of the solution to the single machine scheduling problem for scenario $s = 2$

Example 2: Job Shop Environment

In Fig. 9.7 we give the disjunctive graph representation of the job shop instance. The conjunctive graph of the complete selection that gives the schedule represented in Fig. 9.2 with scenario $s = 1$ is given in Fig. 9.8.

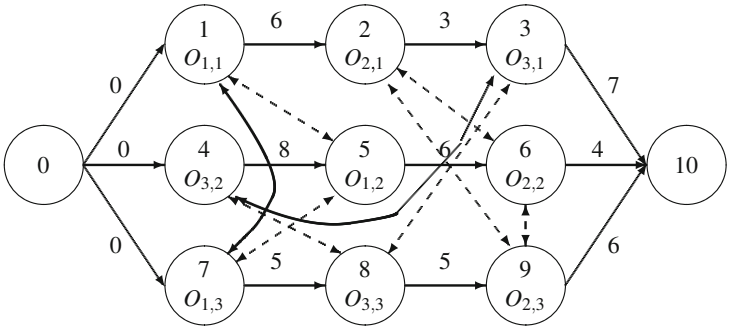


Fig. 9.7: Disjunctive graph corresponding to the instance of the job shop scheduling problem

9.2 Groups of Permutable Jobs: A Solution Structure for Robust Scheduling

This sections present the groups of permutable jobs structure for disjunctive scheduling. Then, it reviews combinatorial optimization problems that have been studied and solved on the group of permutable jobs solution structure.

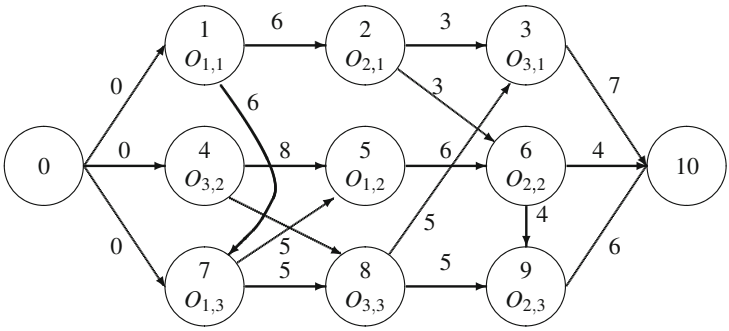


Fig. 9.8: Conjunctive graph corresponding to the solution of the job shop scheduling problem presented for scenario $s = 1$

9.2.1 Groups of Permutable Jobs: A Flexible Solution Representation

A sequence of groups of permutable jobs on a machine M_k is defined as an ordered set partition of the set of jobs $\mathscr{J}$ on machine M_k. An element of each partition is called a group of permutable operations. As for the job sequence, the term "group sequence" is used to name a set of group sequences (one on each machine). This structure was proposed by François Roubellat in the early 1980s [17].

A group sequence $G = (G_i^k)_{i=1,\ldots,v_k}^{k=1,\ldots,m}$ with $\cap_{i=1}^{v_k} G_i^k = \emptyset$ and $\cup_{i=1}^{v_k} G_i^k = \mathscr{J}$ represents a partial job sequence to a disjunctive scheduling problem that specifies on each machine M_k a sequence of v_k groups of permutable operations, such that G_i^k is the i^{th} group on machine M_k and such that all operations inside a group can be permutated without violating the feasibility of the sequence.

Example 1: Single Machine Environment

Let consider the group sequence presented in Fig. 9.9. This sequence is composed of one group composed by jobs $\{J_1, J_3\}$ and of group composed by jobs $\{J_2, J_4, J_5\}$. Of course, the first group starts at time 3 because J_3 cannot start before date 3. The duration of the group is the sum of the durations of the jobs. One can see easily that whatever the order of the jobs in the first group, the jobs of the second group can be executed at time 7 (or earlier if one starts with job J_1). One can also see that the flexibility provided by this group sequence (12 sequences are characterized) has a price, since the makespan is now equal to 17 instead of 14 before.

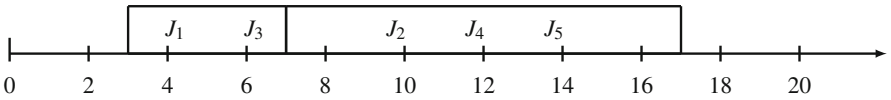


Fig. 9.9: Gantt representation of a group sequence for the single machine problem

Example 2: Job Shop Environment

In Fig. 9.10, two groups of permutable operations are proposed. The first one is composed by the operations of J_1 and J_3 performed on the first machine. The second is composed of the operations of the same jobs on the third machine. One can see that whatever the order of the operations inside each group, the sequence remains feasible. Of course, this flexibility has a price since the makespan is now equal to 32.

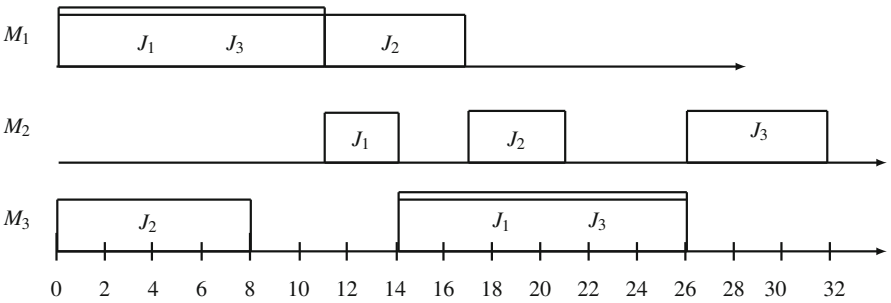


Fig. 9.10: Gantt representation of a group sequence for the job shop scheduling problem

In terms of disjunctive graph, a group sequence matrix corresponds to an incomplete selection that has a particular structure representing the group sequences. Each group is a strongly connected component (via unselected disjunctive arcs) and the selected disjunctive arcs define m totally ordered group sets. We denote by $\mathscr{G}$ the set of (scenario-independent) feasible group sequences.

Given a group sequence $G \in \mathscr{G}$, we denote by $\Sigma(G)$ the set of job sequences that can be obtained from G, and by $\mathscr{C}^s(G)$ the set of schedules issued from the represented job sequences, i.e.:

$$\mathscr{C}^s(G) = \cup_{\sigma \in \Sigma(G)} \mathscr{C}^s(\sigma)$$

For regular objective function, we are interested in the set of earliest schedules that can be issued from a group sequence on a scenario:

$$\mathscr{EC}^s(G) = \cup_{\sigma \in \Sigma(G)} \{EC^s(\sigma)\}$$

It follows that for a group sequence, there is a non unique dominant schedule for a given scenario. Consequently, as a solution representation, a group sequence provides (much) more flexibility. Indeed a group sequence represents $\Pi_{i=1,\ldots,v_k}^{k=1,\ldots,m}|G_i^k|$ different job sequences and earliest schedules.

In turn, there is a much bigger number of feasible group sequences. The number of different weak orders on set $\mathscr{J}$ is equal to

$$b_n = \sum_{k=0}^{n} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^n.$$

Then, for the single machine problem we have $|\mathscr{G}| = b_n$. For the job shop problem, since not all job sequences are feasible, neither are all group sequences and the number of group sequences is bounded by $|\mathscr{G}| \leq mb_n$.

Finally, we have to remark that the set of feasible job sequences maps the set of feasible group sequences such that each group has a single job.

9.2.2 Combinatorial Optimization Problems on Group Sequences

Several combinatorial optimization problems can be defined on the group sequence solution representation and have been studied in the literature. We restrict to the case where earliest schedules are dominant, which corresponds to regular scheduling objective functions. As a given group sequence G represents in general an exponential number of feasible earliest schedules, given a scenario, a question arises on how to select one schedule among the represented ones. In a robust optimization framework, we assume that, given a disjunctive scheduling problem and a scenario set, a group sequence G is computed as a first-stage decision set. In a second stage decision setting, once the scenario is revealed, we assume that one of the represented job sequence (and its corresponding earliest schedule) is selected with algorithm $A(G,s)$. A typical example would be to define $A(G,s)$ as a list scheduling algorithm that selects an order inside each group according to a priority rule. In fact, we can identify the set of list scheduling algorithms compatible with G and the set of job sequences represented by G. This gives rise to several combinatorial optimization problems.

9.2.2.1 Best Earliest Schedule Within a Fixed Group Sequence for a Fixed Scenario

The objective of the second stage algorithm $A(G,s)$ is naturally to obtain the best schedule according to the realized scenario, which yields problem (GP_1). This gives a lower bound on the performance that an algorithm $A(G,s)$ can reach.

$$(GP_1) \quad \min_{\sigma \in \Sigma(G)} L^s_{\max}(\sigma)$$

As we can define a group sequence G such that $|G| = 1$ for the one machine problem, finding the best schedule in this sequence amounts to solve the one-machine problem itself (denoted by $1|r_i|L_{\max}$), which is an NP-hard problem).

Methods that explicitly solve problem (GP_1) for shop problems in an exact or approximated way can be found in [20, 24]. We illustrate these methods for the makespan objective (by considering that all $d_{k,j}$ are equal to 0). These methods rely on the computation for each operation $O_{k,j}$ of the head of the operation, noted $\theta_{k,j}$, i.e., a lower bound for the earliest starting time, and a tail of the operation, noted $\theta'_{k,j}$, i.e., a lower bound of the time between the operation's latest completion time ($C_{k,j}$) and the end of the schedule ($C_{\max}$).

Heads and tails, which are classical notions in shop scheduling, are adapted for groups of permutable operations. The computation of $\theta_{k,j}$ involves the computation of a lower bound for the earliest completion time for each predecessor of operation $O_{k,j}$: the predecessor operation in job routing ($O_{\pi^k_j,j}$) and the predecessor group on the same machine

Let g^k_j the position of the group containing operation $O_{k,j}$ on machine M_k. Then, $G^k_{g^k_j-1}$ (resp. $G^k_{g^k_j+1}$) is its predecessor (resp. successor) group on machine M_k. $\gamma(G)$ denotes a lower bound of the completion time of group G. The computation of $\gamma(G)$ is based on a one-machine relaxation by making the assumption that each machine has an infinite capacity [20, 24]. Below, a lower bound for the earliest completion time of operation $O_{k,j}$ is denoted $\chi_{k,j}$.

For an operation $O_{k,j}$, its head is computed as follows:

$$\theta_{k,j} = \max(r_{k,j}, \gamma(G^k_{g^k_j-1}), \chi_{\pi^k_j,j}) \tag{9.6}$$

$$\gamma(G^k_i) = C_{\max} \text{ of } 1|r_{k,j}|C_{\max} \text{ with } r_{k,j} = \theta_{k,j}, \quad \forall O_{k,j} \in G^k_i$$

$$\chi_{k,j} = \theta_{k,j} + p_{k,j}$$

Because of the symmetry of heads and tails, tails can be computed as heads using a reversed version of Eq. (9.2.2.1): rather than starting the computation at the beginning of the scheduling problem, the computation begins at the end. We use below symmetrical intermediate values $\gamma'(G)$ and $\chi'_{k,j}$ to compute the tail $\theta'_{k,j}$.

Ω^k_j denotes the machine that follows machine M_k in job J_j routing.

$$\theta'_{k,j} = \max(\gamma'(G^k_{g^k_j+1}), \chi'_{\Omega^k_j,j})$$

$$\gamma'(G^k_i) = C_{\max} \text{ of } 1|r_{k,j}|C_{\max} \text{ with } r_{k,j} = \theta'_{k,j}, \quad \forall O_{k,j} \in G^k_i$$

$$\chi'_{k,j} = \theta'_{k,j} + p_{k,j}$$

These heads and tails can be directly used for the computation of a valid lower bound for any regular objective. For example, for the makespan objective, a lower bound is :

$$\max_{k\in\mathscr{M}} \max_{i=1,\ldots,v_k} (\gamma(G_i^k))$$

For the makespan, this lower bound can be improved using the one-machine relaxation proposed by Carlier and Chretienne [10] also based on the computation of heads and tails. In our case the relaxation is made on the groups instead of machines; for each group a lower bound is computed using the exact method of [9]. The maximum value for all groups represents an improved lower bound for the makespan.

In [20], it was shown that the computation of a lower bound for the best earliest schedule has a complexity of $O(n\log n)$ for any regular objective.

9.2.2.2 Worst Earliest Schedule Within a Fixed Group Sequence for a Fixed Scenario

This problem, denoted (GP_2) seeks to determine the worst performance that any algorithm $A(G,s)$ can achieve on a given scenario. This gives an upper bound on the performance of the second stage algorithm. In conjunction with problem (GP_1), we may obtain a lower and an upper bound of the performance of any second stage algorithm compatible with G on a given scenario.

$$(GP_2) \qquad \max_{\sigma\in\Sigma(G)} L^s_{\max}(\sigma)$$

As for the preceding problem, a limit case is to consider a single group of n jobs for the one-machine problem. In this case, we obtain maximization one-machine problems, which are in general easier than their minimization counterpart as shown in [3] that provide polynomial algorithms and complexity proofs for several cases. Furthermore, in [2], it was shown that (GP_2) is polynomial for any disjunctive scheduling problem. The calculation of this worst earliest schedule relies on the computation of the worst earliest starting time and the worst earliest completion time for each operation denoted $\underline{\tau}_{k,j}$ and $\underline{C}_{k,j}$ respectively.

The computation of $\underline{\tau}_{k,j}$ corresponds to executing this operation at its worst position where all its predecessors are placed at their worst latest time. This problem can be formulated as follows :

$$\underline{\tau}_{k,j} = \max(r_{k,j}, \underline{C}_{\pi^k_j,j}, \max_{O_{k,l}\in G^k_{g^k_j-1}} \underline{C}_{k,l}) \tag{9.7}$$

To compute $\underline{C}_{k,j}$, either the worst earliest completion time of operation $O_{k,j}$ does not depend on another operation of the same group, in this case the first term of the following formula is used. Otherwise, in the worst case, the operation ends last in its group and the second term is used.

$$\underline{C}_{k,j} = \max(\underline{\tau}_{k,j} + p_{k,j}, \max_{O_{k,l} \in G^k_{g^k_j}, l \neq j} \underline{\tau}_{k,l} + \sum_{O_{k,l} \in G^k_{g^k_j}} p_{k,l}) \tag{9.8}$$

The computation of $\underline{\tau}_{k,j}$ and $\underline{C}_{k,j}$ can be performed efficiently by longest path computation in a special conjunctive graph, as described in [2]. We do not describe this graph here for sake of conciseness.

Finally, $\max_{J_i \in \mathscr{J}, M_k \in \mathscr{M}} (\underline{C}_{k,j} - d_j)$ represents the worst case evaluation for the groups of permutable operations for the L_{max} objective. More generally, the worst case evaluation can be computed in polynomial time for regular min-max objectives.

This calculation can also be used as an upper bound for min-sum regular objectives.

9.2.2.3 Worst Latest Schedule Within a Fixed Group Sequence for a Fixed Scenario

This problem, denoted (GP_2') is similar to (GP_2) with due dates for the jobs $(d_{k,j})$. It seeks to determine the worst performance that any algorithm $A(G,s)$ can achieve on a given scenario such that there is no late schedule described by the groups of permutable operations.

For this, it needs the computation of the worst latest starting time $\overline{\tau}_{k,j}$ and the worst latest completion time $\overline{C}_{k,j}$ of an operation $O_{k,j}$ which is similar to the computation of $\underline{\tau}_{k,j}$ and $\underline{C}_{k,j}$. A reverse version of Eqs. (9.7) and (9.8) can be used. Rather than starting from the beginning, the computation starts from the end using the due dates of jobs as explained in [19, 23].

$\overline{\tau}^k_j$ and $\overline{C}^k_j$ can be expressed as follows:

$$\overline{C}_{k,j} = \min(d_{k,j}, \overline{\tau}_{\Omega^k_j,j}, \min_{O_{k,l} \in G^k_{g^k_j+1}} \overline{\tau}_{k,l})$$

$$\overline{\tau}_{k,j} = \min(\overline{C}_{k,j} - p_{k,j}, \min_{\forall O_{k,l} \in G^k_{g^k_j}, l \neq j} C_{k,l} - \sum_{\forall O_{k,l} \in G^k_{g^k_j}} p_{k,l}) \tag{9.9}$$

9.2.2.4 Flexibility Maximization with a Bounded Objective for a Fixed Scenario

Without any assumption on the second stage algorithm $A(G,s)$, a question arises to maximize a flexibility measure of the group sequence G, denoted $flex(G)$, while ensuring an upper bound UB on the objective. Intuitively, this allows to propose a group sequence such that the largest number of represented job sequences (or with an alternative view, the largest number of compatible second stage algorithms) satisfy the upper bound.This yields problem (GP_3):

$$(GP_3) \quad \max_{G \in \mathscr{G}} flex(G) \text{ s.t. } L^s_{\max}(\sigma) \leq UB, \forall \sigma \in \Sigma(G)$$

As an outcome, we obtain a performance guarantee on the worst earliest schedule for the considered scenario but we have no indication of the best achievable objective, which would require to solve (GP_1). As problem (GP_1) is generally NP-hard, an alternative from computing the best schedule represented by a group sequence is to ask that the group sequence represents a fixed sequence σ_0, which yields

$$(GP_4) \quad \max_{G \in \mathscr{G}} flex(G) \text{ s.t. } \sigma^0 \in \Sigma(G), L^s_{\max}(\sigma) \leq UB, \forall \sigma \in \Sigma(G)$$

Provided that σ^0 is selected as the optimal solution for scenario s, then solving (GP_4) gives for a given scenario a group sequence G of maximal flexibility such that any job sequence σ issued from G (or any list scheduling algorithm compatible with G) verifies $L^s_{\max}(\sigma^0) \leq L^s_{\max}(\sigma) \leq UB$.

Several flexibility measures are available from the literature. The most natural one is the number of represented job sequences/earliest schedules

$$flex_1(G) = |\Sigma(G)|$$

In [2], a surrogate flexibility measure is used, as the number of groups. Indeed it holds intuitively that generating less groups yield more flexibility. We define this measure as

$$flex_2(G) = |G|$$

To normalize this criterion, for a disjunctive problem with m machines we can define.[1]

$$flex_3(G) = \frac{mn - |G|}{mn - m}$$

In case of full flexibility, we have m groups and $flex_3 = 100\,\%$. In case of no flexibility we have mn groups and $flex_3 = 0\,\%$.

In [2], an $O(n^3)$ algorithm has been proposed to solve problem (GP_3) with $flex_1$ for the one-machine problem without release dates, and in the case where the due dates are agreeable, i.e. for any two jobs $i, j \in \mathscr{J}$, $p_i \leq p_j \Leftrightarrow d_i \leq d_j$. The same algorithms solves also (GP_4) with $flex_1$ for the same problem, but without the restriction of agreeable due dates. In the same context, (GP_3) is solved with $flex_2$ by a simpler $O(n \log n)$ algorithm. Adding now release dates, (GP_4) is solved with $flex_1$ by an $O(n^7)$ algorithm and with $flex_2$ by an $O(n^4)$ algorithm. Hence $flex_2$ yields generally simpler problems than $flex1$.

By varying UB, different compromise solutions can be found between the flexibility criterion and the represented schedule performance. This was illustrated for a job in [2], where a (GP_4) was heuristically solved with $flex_2$. The represented job sequence σ^0 was set to the optimal job shop solution and UB was set to different values above $L^s_{\max}(\sigma_0)$. In [12], the two-machine flow-shop, open-shop and job shop

[1] This was initially proposed in [12] for $m = 2$.

problems are studied. A problem dual to (GP_3) with $flex_2$ is also considered, in the sense that it consists in minimizing the makespan while the number of groups is bounded.

$$(GP_5) \quad \min_{G\in\mathscr{G},\sigma\in\Sigma(G)} L^s_{\max}(\sigma) \text{ s. t. } k_1 \leq |G| \leq k_2$$

For the 2-machine flow shop problem (GP_3) and (GP_5) with $flex_2$ are both NP-hard in the strong sense. However, for any integer $1 \leq k \leq n$ the authors propose a heuristic to build a group sequence G such that $|G| = k$ a job sequence σ such that $\frac{C^s_{\max}(\sigma)}{C^s_{\max}(\sigma^*)} \leq \frac{|G|+1}{|G|}$, where σ^* is the job sequence that minimizes the makespan with no restriction on the number of groups. This is a theoretical bound on the makespan increase when the flexibility is increased. They also propose a heuristic for (GP_4) with $flex_2$. Complexity results are also given for (GP_5) and $flex_2$ for the two machine job shop and flow shop problems.

9.2.2.5 Recoverable Robust Optimization for a Fixed List Scheduling Algorithm on a Scenario Set

If we consider now a fixed list scheduling algorithm $A(G,s)$ that outputs a job sequence that is both compatible with group sequence G and feasible for scenario s, we may seek the group sequence that maximizes the robustness of the earliest schedule selected by the list scheduling algorithm according to the realized scenario. We obtain problem (GP_6).

$$(GP_6) \quad \min_{G\in\mathscr{G}} \max_{s\in S} L^s_{\max}(A(G,s))$$

Note that compared with the job sequence representation, the group sequence representation introduces a third decision level. The first decision level builds the group sequence. The second decision level selects the job sequence. The third decision stage selects the schedule. In Sect. 9.3 we present a method to solve (GP_6) for the single machine problem that we compare with the standard robust scheduling method.

9.3 Solution Methods: A Recoverable Robust Approach Based on Groups of Permutable Operations

Using the concept of recoverable robustness (GP_6) proposed in Sect. 9.2, we present in this section a mixed integer linear program (MILP) and a heuristic method for the maximum lateness minimization on the single machine problem. Given a group sequence and according to the realized scenario, the $A(G,s)$ algorithm schedules the jobs inside a group following the Earliest Release Date (ERD) rule. To evaluate the

interest of the recoverable robust approach, a MILP model as well as a tabu search algorithm are proposed for standard robust scheduling method. Experimental tests are performed and comparisons are given.

9.3.1 MILP Formulation

In order to simulate the ERD rule, a list of predecessors and successors denoted $Prec_j^s$ and $Succ_j^s$ are defined for each job J_j and each scenario s as follows:

$$Prec_j^s = \{J_i \in \mathscr{J} / (r_i^s < r_j^s) \text{ or } (r_i^s = r_j^s \text{ and } i < j)\}$$
$$Succ_j^s = \{J_i \in \mathscr{J} / (r_i^s > r_j^s) \text{ or } (r_i^s = r_j^s \text{ and } i > j)\}$$

We define binary variables $y_{j,q}$ equal to 1 if job J_j is in group G_q (i.e. the group at position q), and 0 otherwise. $C_j^s \geq 0$ is the completion time of job J_j in scenario s.

- Assignment constraints: assign each job to exactly one group

$$\sum_{q=1}^{n} y_{j,q} = 1, \quad \forall j \in \{1,\ldots,n\} \tag{9.10}$$

- Non-overlapping constraints: for a given sequence, the completion time of a job J_j is at least equal to the release date of J_i plus the duration of the jobs that are between J_i and J_j in the sequence, including these two jobs. The following two constraints compute the minimum value of the job completion time under each scenario as set by the list scheduling algorithm. The case were J_i and J_j are in the same group is considered by constraints (9.11), whereas constraints (9.12) consider the case where these two jobs are not in the same group.

$$r_i^s + p_i^s + \sum_{l \in (Succ_i^s \cap Prec_j^s)} p_l^s y_{l,q} + p_j^s - M(2 - y_{i,q} - y_{j,q}) \leq C_j^s,$$
$$\forall i,j,q \in \{1,\ldots,n\}, j \in Succ_i^s, \forall s \in \mathscr{S} \tag{9.11}$$

According to the ERD rule, the expression $\sum_{l \in (Succ_i^s \cap Prec_j^s)} p_l^s y_{l,q}$ computes the total duration of the jobs between J_i and J_j except p_i and p_j.

$$r_i^s + p_i^s + \sum_{l \in Succ_i^s} p_l^s y_{l,q} + \sum_{f=q+1}^{q'-1} \sum_{l=1}^{n} p_l^s y_{l,f} + \sum_{l \in Prec_j^s} p_l^s y_{l,q'} + p_j^s$$
$$-M(2 - y_{i,q} - y_{j,q'}) \leq C_j^s, \quad \forall i,j,q,q' \in \{1,\ldots,n\}, q' > q, \forall s \in \mathscr{S} \tag{9.12}$$

The total duration between the jobs J_i and J_j is represented by the expression $\sum_{l \in Succ_i^s} p_l^s y_{l,q} + \sum_{f=q+1}^{q'-1} \sum_{l=1}^{n} p_l^s y_{l,f} + \sum_{l \in Prec_j^s} p_l^s y_{l,q'}$. Given a scenario s and according to the ERD rule, expressions $\sum_{l \in Succ_i^s} p_l^s y_{l,q}$ and $\sum_{l \in Prec_j^s} p_l^s y_{l,q'}$ com-

pute the total duration of J_i successors and the total duration of J_j predecessors, respectively. The remaining expression $\sum_{f=q+1}^{q'-1} \sum_{l=1}^{n} p_l^s y_{l,f}$ compute the total duration of groups between those including the two jobs.

- The maximum lateness of a schedule

$$L_{\max} \geq C_j^s - d_j^s, \quad \forall j \in \{1,\ldots,n\}, \forall s \in \mathscr{S} \tag{9.13}$$

- Objective function

$$\min L_{\max}$$

This model contains n^2 binary variables, $n|\mathscr{S}|+1$ continuous variables and $O(n^4|\mathscr{S}|)$ constraints.

9.3.2 Tabu Search Algorithms

The proposed MILP can only be used to solve small problem instances. To overcome the difficulty of solving large instances, we propose a tabu search heuristic to solve larger instances. The metaheuristic works as follows. Starting from an initial solution and defining the neighborhood structure, the procedure selects the best appropriate neighbor solution. The selected solution can be chosen if it is not in the tabu list. Otherwise this solution is rejected and the procedure seeks another neighbor solution. The process is repeated if the global stopping condition is not reached.

- *Step 1.* Initial solution: the initial sequence of groups is obtained by sorting the jobs in their due dates increasing order over the first scenario $s = 1$. From this sequence, n groups are created, by assigning each job to one group.
- *Step 2.* Selection of the best neighbor: starting with the current group sequence solution, each neighbor is evaluated and the best non tabu is kept.
- *Step 3.* Stopping condition: the algorithm stops when a global time limit, fixed to 30 s, is reached.

In the following, the implementation of the encoding, neighborhood structure and tabu list for the proposed algorithm are respectively described.

9.3.2.1 Encoding

Because the group sequence and the composition of groups do not depend on the scenario, a solution of the problem can simply be encoded by a vector of size n. Let v be the used vector for encoding a solution and g_j the index of the group to which job J_j is assigned on the solution sequence. A complete solution is encoded by assigning a value g_j to vector v for each j, $1 \leq j \leq n$.

9.3.2.2 Neighborhood Definition

We denote by $|G|$ the number of groups. Four neighborhood are defined as follows:

- *Groups swap:* let v_j the position to swap, select v_k and swap the values v_j and v_k. This neighborhood swaps the groups assigned to jobs J_j and J_k.
- *Group insert:* insert a job J_j of group v_j to an existing group v_k.
- *Group split:* split a group into two groups. The fact that the sequence inside the group depends on the realised scenario makes the split infeasible. To rectify it, we order the jobs inside a group in increasing order according to the average due dates of the jobs over the scenarios.
- *Groups fusion:* merge two consecutive groups into one.

9.3.2.3 Tabu List

The tabu list contains solutions that are recently selected and prevents to choose them again. The experiment analysis do not really show the contribution of a tabu list. Therefore, we have fixed a size of $10n$ for the Tabu list, which gives relatively better results.

9.3.3 Solution Algorithms for the Standard Robust Scheduling Method

To evaluate our algorithms for the recoverable robust approach based on groups of permutable jobs, the algorithms presented above are compared with those of the standard robust scheduling method (without groups of permutable operations). In order to do that, we propose a mixed integer linear program (MILP) and a tabu search heuristic for the standard robust scheduling method. The MILP as well as the tabu search proposed algorithms are briefly presented.

Positional variables $x_{j,k} \in \{0,1\}$ are defined for this model and $L_{\max}$ is a continuous variable to minimize. Variable $x_{j,k}$ takes value 1 if the job J_j is in position k, and 0 otherwise.

- Assignment constraints: assign one job at each position, and one position to each job

$$\sum_{j=1}^{n} x_{j,k} = 1, \quad \forall k \in \{1,\ldots,n\}$$

$$\sum_{k=1}^{n} x_{j,k} = 1, \quad \forall j \in \{1,\ldots,n\}$$

- Guarantee that the worst earliest schedule $L_{\max}$ is larger than the maximum lateness of each scenario

$$L_{\max} \geq \sum_{j=1}^{n} r_j^s x_{j,k} + \sum_{q=k}^{k'} \sum_{j=1}^{n} p_j^s x_{j,q} - \sum_{j=1}^{n} d_j^s x_{j,k'},$$
$$\forall k, k' \in \{1,\ldots,n\}, k \geq k', \forall s \in \mathscr{S}$$

- Objective

$$\min L_{\max}$$

This model contains n^2 binary variables, one continuous variable and $O(n^2|\mathscr{S}|)$ constraints.

To make a fair comparison of the two robust solution paradigms, the proposed tabu search algorithm is based on the same principles as before and starts with the same initial solution. In the following, we describe the implementation of the encoding and the neighborhoods structure, without changing the rest of the algorithm.

A vector v' of size n is used to encode a solution in which v'_j represents the position of job J_j in the schedule. The two used neighborhoods are the following:

- *Position swap:* This neighborhood performs an exchange of the positions of two jobs. Let v_j be the position to swap, select v'_k and swap the values v'_j and v'_k.
- *Position insert:* insert a job at position v'_j to position v'_k. The jobs between v'_j and v'_k will be shifted.

9.3.4 Computational Experiments

The algorithms have been evaluated on randomly generated instances using the following scheme. We first generate (p_j^1, r_j^1, d_j^1) for the scenario $s = 1$ which is called "reference scenario", then for each $s \in \mathscr{S}$, $s \neq 1$, uncertainty was generated from the data (p_j^1, r_j^1, d_j^1). Processing times p_j^1 were uniformly generated in the interval $[1,50]$ and we denote by $P = \sum_{j=1}^{n} p_j^1$ the total processing times on the reference scenario. Release and due dates were generated in the intervals $[1, \mu P]$ and $\left[\alpha - \frac{\beta}{2}P, \alpha + \frac{\beta}{2}P\right]$, respectively. For modelling the uncertainty on the data, for each scenario $s \in \mathscr{S}$, $s \neq 1$, (p_j^s, r_j^s, d_j^s) are generated uniformly from the "reference scenario" by taking values from the intervals $\left[1 - \omega p_j^1, 1 + \omega p_j^1\right]$, $\left[1 - \omega r_j^1, 1 + \omega r_j^1\right]$ and $\left[1 - \omega d_j^1, 1 + \omega d_j^1\right]$, respectively. Parameters μ, α and β take a fixed value $\{0.5\}$, $\{1\}$ and $\{1\}$, respectively. The last parameter ω takes values from the set $\{0.2, 0.4, 0.6\}$. For each couple (n,s) ten instances are generated in which $n \in \{10, 25, 40, 100\}$ and $\mathscr{S} \in \{2, 5, 10\}$. The experiments have been run for the 360 instances on an Intel i7-4770 CPU 3.40 GHz computer with 8 GB.

We evaluated and compared the performance of the solutions obtained by the MILP models and the tabu search algorithms. We call *RRA* the recoverable robust approach and *SRA* the standard robust approach. Let $z_E(A)$ and $z_H(A)$ denote the objective value of the algorithm $A \in \{RRA, SRA\}$ returned respectively by the exacts methods given by MILPs and the heuristic one given by the tabu search algorithm.

In Table 9.5, we provide the aggregate results for each n. The statistics take account of the average CPU times ttb needed to reach the best solution for both tabu search algorithms. The average gap between the exact solution found and the tabu search solution are given by gap columns. gap takes value $(z_H(SRA) - z_E(SRA))/z_H(SRA)$ in the SRA case and $(z_H(RRA) - z_E(RRA))/z_H(RRA)$ in the RRA one. The last column $\Delta = (z_H(SRA) - z_H(RRA))/z_H(SRA)$ gives the average gap between the two tabu search algorithms.

Table 9.5: Experimental comparisons

	SRA		RRA		
n	$ttb(s)$	gap	$ttb(s)$	gap	Δ
10	0.01	1.7 %	0.01	0 %	24.79 %
25	0.14	2.3 %	0.15	–	17.28 %
40	0.39	–	0.52	–	16.83 %
100	3.21	–	12.58	–	13.71 %

The experiments show the limits of MILP models, especially for the recoverable robust approach. In ttb columns, one can see that the RRA tabu search algorithm has a faster growth than the one for SRA. However, the difference remains not very significant and both algorithms have comparable CPU times for considered instances. As expected finally, column $Delta$ reveals the benefits of the recoverable robust approach to obtain better worst case maximum lateness values. This is due to the fact that the RRA can react to the realized scenario thanks to the group structure.

9.4 Using Groups of Permutable Operations in an Industrial Context

As mentioned in Sect. 9.2.2 robust machine scheduling based on groups of permutable operations is composed of three decisions stages. The second stage decision set consists, once the scenario is revealed, in selecting a job sequence and we assume in this chapter that the earliest schedule is selected from a given job sequence. In an industrial context the selection of a job sequence has to be made in a very short time. In order to fulfill this timing constraint, either heuristics adapted to groups of permutable operations may be used (as the ERD list scheduling algorithm used in the previous section) or the selection may be done by a human operator during the execution of the schedule. This section describes several alternative ways of selecting job sequence from a group sequence on a realized scenario (index s is consequently dropped), keeping industrial requirements in mind.

9.4.1 Heuristics for the Reactive Phase of Groups of Permutable Operations

To solve the job shop scheduling problem one of the most common approaches is the use of heuristics based on priority dispatching rules (PDR), which are rules used to select the next job to process from jobs awaiting service on a resource. The lower bounds presented in Sect. 9.2.2 are used to build PDRs for groups of permutable operations when the objective is to minimize the makespan.

First a PDR based on the operation's tail is proposed. The idea is to give more priority to an operation which exhibits a large tail. As the tails of different operations may be equal, the rule Shortest Processing Time (SPT), which aims at selecting the operation with the shortest imminent processing time, is used to break the ties. This heuristic named SQUTAIL (Square Tail) is formulated as:

$$\min(p_{k,j} - \theta'_{k,j})$$

From the lower bounds another PDR is proposed, with the following behavior:

1. For each operation waiting in the queue, a partial group of operations is generated;
2. The lower bound for the makespan is computed for these partial generated schedules;
3. The operation with the lowest lower bound is then chosen.

To break the ties, this rule, named LB (Lower Bound), is combined with either rule SQUTAIL or the PDR Most Work Remaining (MWR) which selects the operation with the highest remaining processing time. The use of MWR together with LB is named LB+MWR and the use of SQUTAIL together with LB is named LB+SQUTAIL.

Another heuristic, which is not based on the lower bounds, is also proposed. Shifting Bottleneck (SB) heuristic, described in [1], is a very effective heuristic in job shop scheduling for the makespan. In order to adapt the shifting bottleneck for groups of permutable operations, the relaxation in one-machine problem, is used. In our case, the algorithm is not applied to the machines but to the groups. As the number of re-optimizations is higher than with the classical SB, better performances are expected as well as higher computation time. Another benefit of the relaxation applied to the groups is that all the computed schedules are feasible contrarily to the classical SB which may give schedules that are not feasible. This heuristic is named SB.

In [19], these heuristics have been evaluated on a well-known benchmark for job shop scheduling, the Lawrence's instances composed of 40 instances of 8 different sizes [16]. They show that these heuristics are very effective to evaluate the makespan with the following ranking:

- SQUTAIL is the less efficient (in average a deviation of 13.7 % from the optimal) but the fastest (less than 0.13 s in average with a maximum of 0.46 s),

- LB+MWR and LB+SQUTAIL are quite similar in performance (an average of 3.6 % from the optimal) and computing time (an average of 0.94 s with a maximum of 3.8 s),
- SB is the most effective, giving the optimal for 17 instances (in average a deviation of 1.5 % from the optimal and a maximum of 3.2 %). It is also more time-consuming (an average of 3.7 s with a maximum of 10.74 s).

Each heuristic has specific strengths and weaknesses. To give in a very short time a solution, SQUTAIL is a good compromise, SB is a very effective heuristic in regards with the performance, LB+MWR and LB+SQUTAIL are in between.

9.4.2 A Multi-Criteria Decision Support System (DSS) for Groups of Permutable Operations

Another approach to select a job sequence is to let a human operator choose in real-time the next operation to process within a group of permutable operations, according to the operator's knowledge of the context. In order to make his choice, the operator needs criteria adapted to groups of permutable operations. In Sect. 9.2.2 a best earliest schedule and a worst earliest schedule evaluation for any regular objective within a group sequence for a fixed scenario have been presented. In [23] an adaptation of the free margin to the groups of permutable operations is presented. The so-called free sequential margins allows to evaluate during the execution of the schedule, the schedule lateness.

9.4.2.1 Free Sequential Margin

The free sequential margin computes for an operation according to its earliest execution, the maximum tardiness which ensures that all schedules enumerated in the group sequence will present no tardiness.[2] The free sequential margin of an operation $m_{\text{seq}}(O_{k,j})$ has two components:

- the operation's net margin $m_{\text{sn}}(O_{k,j})$, which is related to the operation itself regardless the other operations of the group.
- the operation's group margin $m_{\text{sg}}(O_{k,j})$, which is related to the other operations of the group.

The computation of the proper free sequential margin of $O_{k,j}$ corresponds to the difference between its worst latest starting time $\overline{\tau}_{k,j}$ (9.9) and its worst earliest starting time $\underline{\tau}_{k,j}$ (9.7).

Using the worst case earliest starting time and the worst case latest completion time, the free sequential margin can be expressed as follows :

[2] Positive lateness.

$$\begin{cases} m_{\text{seq}}(O_{k,j}) = \min(m_{\text{sn}}(O_{k,j}), m_{\text{sg}}(O_{k,j})) \\ m_{\text{sn}}(O_{k,j}) = \overline{C}_{k,j} - p_{k,j} - \underline{\tau}_{k,j} \\ m_{\text{sg}}(O_{k,j}) = \min_{\forall O_{k,l} \in G^k_{g^k_j}, l \neq j} \overline{C}_{k,j} - \sum_{\forall O_{k,l} \in G^k_{g^k_j}} p_{k,l} - \underline{\tau}_{k,j} \end{cases}$$

For a given group of permutable operations, several situations may occur:

- All the free sequential margins of the current group are positive or zero, in that case whatever the chosen operation, all possible permutations on the group will give schedules with no tardiness. Nevertheless, choosing the operation with the highest group margin permits to maximize the margins.
- There is at least one operation in the group which presents negative free sequential margin. In that case, there may be sequences in this group which give tardiness, but it is also possible to have sequences with no tardiness:
- If all net margins are positive then there may be sequences on this group with no tardiness. It is recommended to execute the operation with the highest group margin in order to increase the negative margins, trying to make them become positive.
- If there is at least one operation with a negative net margin then all possible permutations on the group will give late schedules.

An industrial manufacturing scheduling software named ORDO have been developed in France, based on the concept of groups of permutable jobs and using the free sequential margin indicator at the shop floor. At the early 2000s ORDO was used in more than 70 make-to-order manufacturing companies. The software is described in [21] and more references can be found in the book [17].

9.4.2.2 Multi-Criteria DSS

The free sequential margin is the only criterion used in an industrial context (ORDO software) to help the operator choose an operation within a group of permutable operations. However, with only one criterion at his disposal, the operator has little choice to make his decision.

An experiment conducted at the University of Nantes has tried to evaluate if a DSS composed of several criteria could be more efficient [18]. This experiment was realized on a real manufacturing system that can be represented by a six machines job shop problem. During the first stage decision set, groups of permutable operations with a fixed scenario were computed. In the second stage, 18 students at the end of their bachelor degree in production management studies have played the role of the operator. Each student was asked to schedule in real time a single workstation (the same for each student), by choosing an operation within a group. The objective given to the students was to minimize the tardiness, measured by the L_{max}.

The students were separated in two equal groups. The first one has only the free sequential margin at his disposal while the second group has five different criteria to make his choice:

- The best earliest schedule evaluation, which gives the best predictable quality of the schedule if the operation is chosen
- The worst earliest schedule evaluation, which gives the worst predictable quality of the schedule if the operation is chosen
- The operation's free sequential margin,
- The operation's sequence in the routing,
- The operation's processing time

For the second group, students have to explicitly query for each criterion, one by one. Thus, the criteria used to help the operator for taking his decision are registered.

The performance of the human-machine system is measured through the quality of the decision process and not through the scheduling performance. Indeed, it would only take one "bad" decision to downgrade the scheduling performance. The quality of the decision making process is evaluated through the proportion of "good solutions" taken by the operator. A solution is considered "good" if it is not dominated by another potential choice considering the L_{max}.

The results, presented in form of Boxplots (Fig. 9.11), show that with a multi-criteria DSS, the proportion of good solutions increases, and this effect is significant. Using the multi-criteria system the mean proportion of good solutions is 0.86 while it is 0.78 using only the free sequential margin. However, Fig. 9.12 shows that the free sequential margin remains the most used criterion, the best and the worst case evaluation are the less used.

This experiment indicates that with a multi-criteria DSS the quality of the decision process is better. Concerning the criteria, the free sequential margin remains the dominating criterion. This is not surprising because the instruction was to minimize the tardiness and this criterion measures the capacity to absorb expected delays. The best and the worst earliest schedules are less used. This can be explained by the fact that they have a great anticipation effect contrarily to the operation's sequence and the operation's processing time. These two criteria give direct information on the operation and thus are better understood by the operator.

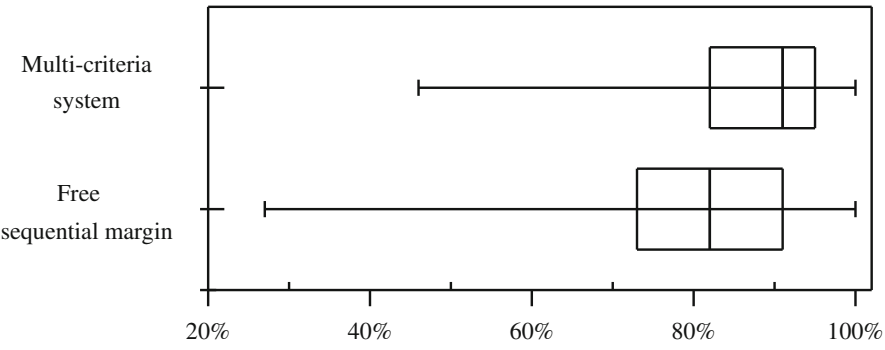


Fig. 9.11: Proportion of good solutions

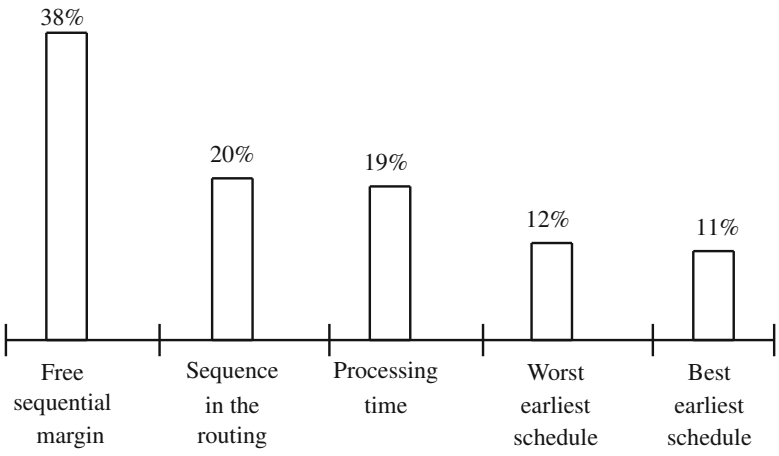


Fig. 9.12: Average proportion of queries by criterion

References

1. Adams, J., Balas, E., Zawack, D.: The shifting bottleneck procedure for job shop scheduling. Manag. Sci. **34**(3), 391–401 (1988)
2. Artigues, C., Billaut, J.C., Esswein, C.: Maximization of solution flexibility for robust shop scheduling. Eur. J. Oper. Res. **165**(2), 314–328 (2005)
3. Aloulou, M., Kovalyov, M., Portmann, M.C.: Maximization in single machine scheduling. Ann. Oper. Res. **129**, 21–32 (2004)
4. Berlinska, J., Drozdowski, M.: Scheduling divisible MapReduce computations. J. Parallel Distrib. Comput. **71**(3), 450–459 (2011)
5. Berlinska, J., Drozdowski, M.: Scheduling multilayer divisible computations. RAIRO Oper. Res. **49**(2), 339–368 (2015)
6. Billaut, J.-C., Moukrim, A., Sanlaville, E. (eds.): Flexibility and Robustness in Scheduling. Wiley, New York (2008)
7. Blazewicz, J., Ecker, K.H., Pesch, E., Schmidt, G., Weglarz, J.: Handbook on Scheduling. From Theory to Applications. Springer, Berlin (2015)
8. Brucker, P.: Scheduling Algorithms, 5th edn. Springer, Berlin (2015)
9. Carlier, J.: The one machine problem. Eur. J. Oper. Res. **11**(1), 42–47 (1982)
10. Carlier, J., Chretienne, P.: An algorithm for solving the job-shop problem. Manag. Sci. **35**(2), 164–176 (1989)
11. D'Ariano A., Pacciarelli, D., Pranzo M.: A branch and bound algorithm for scheduling trains in a railway network. Eur. J. Oper. Res. **183**(2), 643–657 (2007)
12. Esswein, C., Billaut, J.C., Strusevich, V.A.: Two-machine shop scheduling: compromise between flexibility and makespan value. Eur. J. Oper. Res. **167**(3), 796–809 (2005)
13. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W.H. Freeman & Co., New York (1979)
14. Kergosien, Y., Lenté, C., Billaut, J.-C., Perrin, S.: Metaheuristic algorithms for solving two interconnected vehicle routing problems in a hospital complex. Comput. Oper. Res. **40**(10), 2508–2518 (2013)
15. Kouvelis, P., Yu, G.: Robust Discrete Optimization and Its Applications. Kluwer Academic Publishers, Dordrecht (1997)
16. Lawrence, S.: Resource Constrained Project Scheduling: An Experimental Investigation of Heuristic Scheduling Techniques (Supplement). Carnegie-Mellon University, Pittsburgh (1984)

17. Lopez, P., Roubellat, F.: Production Scheduling. ISTE/Wiley, London/Hoboken (2008)
18. Mebarki, N., Cardin, O., Guérin, C.: Evaluation of a new human-machine decision support system for group scheduling. In: 12th IFAC, IFIP, IFORS, IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, vol. 12, no. 1, pp. 211–217 (2013)
19. Pinot, G.: Coopération homme-machine pour l'ordonnancement sous incertitudes. Ph.D. thesis, Nantes University, France (2008)
20. Pinot, G., Mebarki, N.: An exact method for the best case in a group sequence: application to a general shop problem. In: 13th IFAC Symposium on Information Control Problems in Manufacturing, pp. 1269–1274 (2009)
21. Roubellat, F., Billaut, J.-C., Villaumié, M.: Ordonnancement d'atelier en temps réel: d'ORABAID à ORDO. Revue d'automatique et de productique appliquées **8**(5), 683–713 (1995)
22. Roy, B., Sussman, B.: Les problemes d'ordonnancement avec contraintes disjonctives, Note D.S. No. 9 bis, SEMA, Paris (1964)
23. Thomas, V.: Aide à la décision pour l'ordonnancement d'atelier en temps réel. Ph.D. thesis, Paul Sabatier University, France (1980)
24. Yahouni, Z., Mebarki, N., Sari, Z.: New lower bounds for the best-case schedule in groups of permutable operations. IFAC-PapersOnLine **48**(3), 1134–1139 (2015)

Chapter 10
How Robust is a Robust Policy? Comparing Alternative Robustness Metrics for Robust Decision-Making

Jan H. Kwakkel, Sibel Eker, and Erik Pruyt

Abstract Nowadays, decision-makers face deep uncertainties from a myriad of external factors such as climate change, population growth, new technologies, and economic developments. The challenge is to develop robust policies, which perform well across all possible resolutions of the uncertainties. One approach for achieving this is to design a policy to be adapted over time in response to how the future actually unfolds. A key determinant for the efficacy of such an adaptive policy is the specification of when and how to adapt it. This specification depends on how robustness is being operationalized. To date, there is little guidance for selecting an appropriate robustness metric. In this chapter we address this problem, using a case study of designing a policy for stimulating the transition of the European energy system towards more sustainable functioning using five different robustness metrics. We compare the policies as identified by each metric and discuss their relative merits. We highlight that the different robustness metrics emphasize different aspects of what makes a policy robust. More specifically, measures that separate dispersion and the mean, effectively doubling the number of objectives, provide very valuable information on the trade-offs between the mean performance of the policy and dispersion around this mean. We also discuss, based on our case, why analysts should use multiple robustness metrics.

J.H. Kwakkel (✉) • S. Eker • E. Pruyt
Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands
e-mail: J.H.Kwakkel@tudelft.nl; s.eker@tudelft.nl; e.pruyt@tudelft.nl

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_10

10.1 Introduction

In many planning problems, planners face major challenges in coping with uncertain and changing physical conditions, and rapid unpredictable socio-economic development. How should society prepare itself for this confluence of uncertainty? Given the presence of irreducible uncertainties there is no straightforward answer to this question. Effective decisions must be made under unavoidable uncertainty [10, 23]. The acceptance of uncertainty as an inevitable part of long-term decision-making has given rise to the development of new tools and approaches (see Walker et al. [36] for a review).

Robust decision making is a paradigm example of the new approach to model-based decision support in the face of deep uncertainty [13, 22]. In this approach, a very large ensemble of plausible futures spanning the various key uncertain factors is created [20]. This ensemble serves as a test bed for candidate policies. Through scenario discovery, the key vulnerabilities and opportunities of a candidate policy are identified [6]. In light of this, an iterative process of (re)design of candidate policies takes place, aimed at improving the overall robustness of the policy.

Typically, the iterative redesign of candidate policies involves the inclusion of actions whose implementation is conditional on how the future unfolds [13]. The challenge here is to avoid implementing these actions either too early or too late. Very recently, robust multi-objective optimization has been suggested as a technique for supporting the search for finding the right conditions [14]. However, a variety of alternative operationalizations of robustness have been used within a multi-objective optimization framework (e.g., [13, 15, 17, 21, 22]). In these various operationalizations, robustness is understood either as reducing the uncertainty about the expected consequences of a given policy. So no matter how the future plays out, the policy performance falls in a narrow bandwidth. Or, alternatively, robustness is understood as minimizing the undesirable outcomes. So, no matter how the future unfolds, policy performance will be satisfactory. This raises the question how the choice of the robustness metric affects the final design of an adaptive policy or plan. It also raises the question whether some robustness metrics always outperform other robustness metrics. Therefore, insight into the consequences of different robustness metrics can help analysts in choosing a (set of) metric(s) that is appropriate for the case at hand, and improve awareness regarding the relative merit of alternative robustness metrics.

In this chapter, we apply five different robustness metrics to the same case, allowing us to compare the results and providing insight into the relative merits of each of these five metrics. We start from the European energy transition case studied by Hamarat et al. [14]. This case focuses on finding an adaptive plan, built on the European emission-trading scheme that maximizes the potential of achieving the emission reduction targets set out by the European Commission. A variety of actions can be taken to help in achieving the emission reduction targets. The key question is to identify the conditions under which a given action should be taken. This can be formulated as a robust multi-objective optimization problem. In this chapter, we explore the consequences of alternative robustness metrics using this case.

The chapter is structured accordingly. In Sect. 10.2 we introduce the alternative robustness metrics. Section 10.3 provides additional details on the multi-objective optimization formulation and the approach for solving this problem, and briefly introduces the simulation model of the European energy system, the key uncertain factors that need to be accounted for, and the case specific formulation of the robust multi-objective optimization problem. Section 10.4 contains the results for each of the five robustness metrics and their comparison. A discussion of the results and their implications is presented in Sect. 10.5.

10.2 Measuring Robustness

Robust optimization methods aim at finding, in the presence of uncertainty about inputs, optimal outcomes that are not overly sensitive to any specific realization of the uncertainties [1, 3–5, 19]. In robust optimization, the uncertainty that exists about the outcomes of interest is described through a set of scenarios [25]. Robustness is then defined over this set of scenarios. This definition differs from worst-case formulations such as minimax, which can produce very costly and conservative solutions [25].

The way in which robustness is defined over the scenario set can affect the solutions that are being found. A careful choice of the robustness metric(s) is thus paramount. Three families of metrics are available: regret, satisficing, and statistical. Regret based metrics are comparative in character. They compare the performance of a policy option in a given scenario with some performance measure in the same scenario. Regret-based metrics originate from Savage [30]. Savage defines regret as the difference between a given policy's performance in a specific possible future and the performance of the best performing policy option in that specific possible future. A robust policy is the one that minimizes the maximum regret across all alternative possible futures. Alternative regret metrics use some type of baseline performance for a given scenario instead of best performing option [17, 22, 26]. In this chapter, we will not further consider regret-based metrics, for they are very hard to embed efficiently within an optimization routine.

Satisficing metrics aim at maximizing the number of scenarios, which meet a minimum performance threshold. They thus rely on the upfront specification of this performance threshold. A well-known example of this is the domain criterion [31, 32], which focuses on the fraction of the space where a given performance threshold is met; the larger this space, the more robust the policy. Often, this is simplified to looking at the fraction of scenarios, rather than the volume of the space. Recently, the domain criterion made a resurgence under the label of info-gap decision theory [2]. Because of their reliance on a user specified performance threshold, satisficing criteria introduce a new source of uncertainty into the analysis. In part because of this, we ignore these criteria in this chapter.

The third family of robustness metrics are statistical or density based. That is, they look at the distributional character of the outcomes of interest. The basic intuition of these metrics is that a robust policy leaves little uncertainty about the

expected outcomes. That is, the more peaked the distribution of expected outcomes, or the more skewed towards the desired region of outcomes, the more robust the policy. In this chapter, we focus our analysis on this family of metrics and we highlight several exemplary metrics.

For comparative purposes, we include the conservative maximin (or minimax) criterion as our first metric. It is the most conservative and focuses only on the worst case and making this worst case as good as possible

$$f_i(x) = \begin{cases} \max(\mathbf{x}_i), & \text{minimization} \\ \min(\mathbf{x}_i), & \text{maximization} \end{cases} \tag{10.1}$$

where $\mathbf{x}_i$ is a vector of the performance of outcome indicator i in each scenario. In case of minimization, the worst case is the maximum performance. In case of maximization, the worst case is the minimum performance. Note that we take the maximum or minimum over the performance across the set of scenarios, rather than use an optimization for finding the globally worst case. Since this metric only considers the worst case, it is expected to result in conservative and costly solutions. The second metric is based on the intuition that a robust solution will have a good average result with very limited dispersion around it. In mathematical form

$$f_i(x) = \begin{cases} (\mu_i + 1)(\sigma_i + 1), & \text{minimization} \\ (\mu_i + 1)/(\sigma_i + 1), & \text{maximization} \end{cases} \tag{10.2}$$

where μ_i is the mean over the set of scenarios for outcome indicator i and σ_i is the standard deviation. The +1 is included to handle situations where either μ_i or σ_i is close to zero. This metric is essentially a signal to noise ratio, or a form of risk discounting. There are three downsides to this first metric. First, it does not provide insight into the trade-off between improving the mean and reducing the standard deviation. Second, functions that combine the mean and variance are not always monotonically increasing [28]. Third, by using the standard deviation, good and bad deviations from the mean are treated equally [34]. In many cases, however, robustness is understood as minimizing the undesirable deviations from the mean. These problems can be solved in various ways. In this chapter we consider three alternative approaches. Common to these approaches is that they have separate objectives for the mean and some measure of the deviation from this, effectively doubling the number of objectives in the eventual optimization problem.

The third metric is a variant of the approach used by Takriti and Ahmed [34], where we measure the mean and the undesirable deviations away from some target value as separate objectives

$$f_i(x) = \begin{cases} -\mu_i, \sum_{k=1}^{k} (x_k - q_{50})^2 [x_k > q_{50}], & \text{minimization} \\ \mu_i, -\sum_{k=1}^{k} (x_k - q_{50})^2 [x_k < q_{50}], & \text{maximization} \end{cases} \tag{10.3}$$

where q_{50} is the median performance, k is a scenario, x_k is the score for the i-th outcome indicator in scenario k, and the sum is only taken over the cases that meet the specified condition. So, we are taking the sum of squared differences from the median in the undesirable direction. Effectively, this metric uses the sum of squared differences as proxy for the skewness of the distribution.

The fourth metric offers an alternative way of measuring the skewness of the distribution of outcomes, using a quantile-based definition following Voudouris et al. [35]. This quantile-based definition is motivated by the potential unreliability of moment-based definitions of skewness if the density estimate is fat-tailed [8], i.e. if there are many scenarios (data points) in the tail regions.

$$f_i(x) = \begin{cases} -\mu_i, \dfrac{(q_{90}+q_{10})/2 - q_{50}}{(q_{90}-q_{10})/2}, & \text{minimization} \\ \mu_i, \dfrac{(q_{90}+q_{10})/2 - q_{50}}{(q_{90}-q_{10}/2}, & \text{maximization} \end{cases} \tag{10.4}$$

where q_{10}, q_{50}, and q_{90}, are the 10th, 50th, and 90th quantile respectively or the distribution of outcomes for outcome indicator i. The more positive the values of this metric, the more skewed the density estimate is towards the right (higher values). The more negative this value is, the more skewed the distribution is towards the left (lower values). Therefore, for outcome indicators to be maximized, the policy option with higher skewness is preferred, whereas for outcome indicators to be minimized, lower skewness is preferred.

Instead of considering the skewness of the distribution, one can consider the peakedness. Kurtosis is a well-known measure, established originally to describe peakedness. Among several formulations of kurtosis [16], following Voudouris et al. [35], a simple quantile-based metric is adopted here

$$f_i(x) = \begin{cases} -\mu_i, \dfrac{q_{90}-q_{10}}{q_{75}-q_{25}}, & \text{minimization} \\ \mu_i, \dfrac{q_{90}-q_{10}}{q_{75}-q_{25}}, & \text{maximization} \end{cases} \tag{10.5}$$

where q_{10}, q_{25}, q_{75}, and q_{90}, are the 10th, 25th, 75th, and 90th quantile, respectively, of the distribution of outcomes for outcome indicator i. The higher this metric, i.e. the smaller the inter-quartile range compared to the interval between 90th and 10th quantiles, the more peaked the density estimate is around the mean.

10.3 Case

The European Union (EU) has targets for the reduction in carbon emissions and the share of renewable technologies in the total energy production by 2020 [7]. The main aim is to reach 20 % reduction in carbon emission levels compared to 1990

levels and to increase the share of renewables to at least 20 % by 2020. In order to meet the 2020 goals, the EU adopted the European Emissions Trading Scheme (ETS) for limiting the carbon emissions [7]. ETS imposes a cap-and-trade principle that sets a cap on the allowed greenhouse gas emissions and an option to trade allowances for emissions. However, current emissions and shares of renewables show a fragile progress of reaching the 2020 targets. Moreover, the energy system includes various uncertainties related to e.g. technology lifetimes, economic growth, costs, learning curves, and investment preferences. Whether the policy will achieve its targets is at least partly contingent on how these various uncertainties play out.

10.3.1 Model

In this study, a System Dynamics [11, 27, 33] model is used for simulating plausible futures of the EU electricity system. The model represents the power sector in the EU and includes congestion on interconnection lines by distinguishing seven different regions in the EU. These are United Kingdom, Ireland, Italy, and northwest, northeast, middle, southwest, and southeast Europe. Nine power generation technologies are included. These are: wind, PV solar, solid biomass, coal, natural gas, nuclear energy, natural gas with Carbon Capture and storage (CCS), coal gasification with CCS, and large scale hydro power. The model includes endogenous mechanisms and processes related to the competition between technology investments, market supply-demand dynamics, cost mechanisms, and interconnection capacity dynamics.

Figure 10.1 shows the main sub-models that constitute this model at an aggregate level. These are installed capacity, electricity demand, electricity price, profitability and levelised costs of electricity. At an aggregated level, there are two main factors that drive new capacity investments: electricity demand and expected profitability. An increase of the electricity demand leads to an increase in the installed capacity, which will affect the electricity price. This will cause a rising demand, in turn resulting in more installed capacity. On the other hand, decreasing electricity prices will lead to lower profitability and less installed capacity, which will result in electricity price increases. Each sub-model has more detailed interactions within itself and with the other sub-models and exogenous variables and these causal relationships drive the main dynamics of the EU electricity system. More detail on the model can be found in Loonen [24], including a detailed description of all equations and variables.

We are interested in exploring and analyzing the influence of a set of deeply uncertain input variables on the key output variables. In order to explore this uncertainty space, not only parametric but also structural uncertainties are included. For exploring structural uncertainties, several alternative model formulations have been specified and a switch mechanism is used for switching between these alternative formulations. Parametric uncertainties are explored over pre-defined ranges. Table 10.1 provides an overview of the uncertainties, 46 in total.

In many scenarios, the ETS policy alone will not be sufficient to achieve the stated CO_2 emission reduction targets. It thus needs not be complemented with additional actions. The reasons for not achieving the reduction targets differ from one

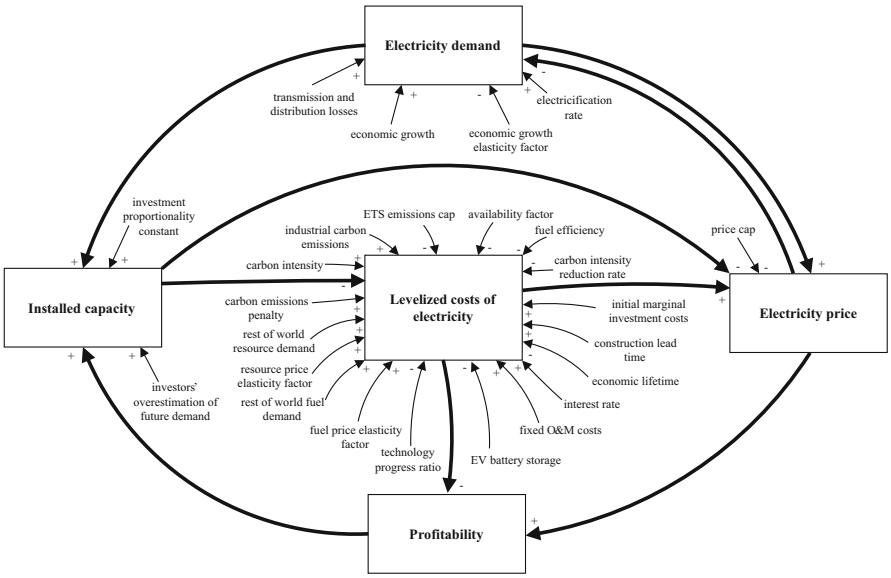


Fig. 10.1: The main causal loops in the EU energy model

scenario to the next. Therefore, the additional actions are scenario dependent. We consider three possible additional actions:

1. Obligatory phase out of older energy generation technology.
2. Subsidy for sustainable energy generation technology. The amount of the subsidy is conditioned on the differences in marginal costs between sustainable and non-sustainable technology.
3. Obligatory decommissioning of non renewable technologies to maintain the achieved fraction of renewable technologies.

Each of these actions is good only in some scenarios. For example, the third action is aimed at scenarios where the use of renewable energy collapses after a quick uptake. To address this, the challenge is to specify up front the conditions under which each of these actions should be implemented. We address this problem using multi-objective robust optimization.

Table 10.1: Specification of the uncertainties to be explored

Name	Description
Economic lifetime	For each technology, the average lifetime is not known precisely. Different ranges for the economic lifetimes are explored for each technology.
Learning curve	For each technology, the extent to which costs will decrease with increasing experience is uncertain. Different progress ratios are explored for each technology.
Economic growth	It is deeply uncertain how the economy will develop over time. Six possible developments of economic growth behaviors are considered.
Electrification rate	The rate of electrification of the economy is explored by means of six different electrification trends.
Physical limits	The effect of physical limits on the penetration rate of a technology is unknown. Two different behaviors are considered.
Preference weights	Investor perspectives on technology investments are treated as being deeply uncertain. Growth potential, technological familiarity, marginal investment costs and carbon abatement are possible decision criteria.
Battery storage	For wind and PV solar, the availability of (battery) storage is difficult to predict. A parametric range is explored for this uncertainty.
Time of nuclear ban	A forced ban for nuclear energy in many EU countries is expected between 2013 and 2050. The time of the nuclear ban is varied between 2013 and 2050.
Price demand elasticity	A parametric range is considered for price demand elasticity factors.

10.3.2 Formulating the Problem

The general optimization problem we are solving is

$$\text{minimize} \quad F(L) = [f_{costs}, -f_{renewables}, -f_{reduction}]$$

where $L = [l_{df}, l_{ad}, l_{sf}, l_{sd}, l_{pr}, l_{dcf}, l_{fth}, l_{tr}]$

$$\begin{aligned} \text{subject to } & 0.5 \leq c_{df} \leq 1, \\ & 0.0 \leq c_{ad} \leq 0.75 \\ & 0.0 \leq c_{sf} \leq 0.5 \\ & 0.0 \leq c_{sd} \leq 20.0 \\ & 1.0 \leq c_{pr} \leq 2.0 \\ & 0.0 \leq c_{dcf} \leq 0.5 \\ & 0.0 \leq c_{fth} \leq 1.0 \\ & 10 \leq c_{tr} \leq 40 \end{aligned}$$

Table 10.2 offers an explanation of each policy leaver l, and the meaning of the subscripts for both l_i and C_i. f_{costs}, $f_{renewables}$, and $f_{reduction}$ are the costs of the policy, the fraction of renewables at the end of the simulation, and the reduction

of emissions of green house gases respectively. The constraints c_i are taken from Hamarat et al. [14] and are based on common sense and case specific considerations.

Various approaches exist for solving multi-objective optimization problems. Over the last decade, substantial advances have been made through the use of genetic algorithms (GA). GA use a population of solutions, which are evolved over the course of the run of the algorithm. This population can be evolved in such a way that it maintains diversity, while continuallymoving towards the Pareto frontier. In this

Table 10.2: List of triggers and their descriptions

	Trigger	Brief description
Action 1	Desired fraction (df)	Desired fraction of renewable technologies.
	Additional decommissioning (ad)	Additional fraction of non-renewable technologies to be decommissioned.
Action 2	Subsidy factor (sf)	Additional fraction of subsidy for renewables.
	Subsidy duration (sd)	Duration for how long the subsidy for the renewables will be active.
	Proximity (pr)	Proximity of cost to the cost of the most expensive non-renewable technology.
Action 3	Decommissioning factor (dcf)	Fraction to be decommissioned for non-renewables when the gap between desired and forecasted fraction for renewables is above the Trigger.
	Forecast time horizon (fth)	Time horizon over which the forecast for the level of renewable fraction is done.
	Trigger (tr)	Proximity of the forecasted renewable fraction to the desired fraction.

way, multiple Pareto front solutions can be found in a single run of the algorithm [9]. Currently, a wide variety of alternative multi-objective evolutionary algorithms are available for solving multi-objective optimization problems [12, 18]. In this study, we use Borg, a state of the art GA where the evolutionary operators co-evolve with the search [12]. In various comparisons, Borg has been demonstrated to be among the best available genetic algorithms for solving multi objective optimization problems [12, 29].

10.4 Results

Figure 10.2 shows the ε-progress of the genetic algorithm for each of the five alternative robustness metrics (10.1)–(10.5). ε-Progress measures how often the genetic algorithm has been able to find a substantially better solution. Over the course of the optimization, this will occur less frequently. As a result, ε-progress stabilizes, indicating that the algorithm has converged. The number of function evaluations required to achieve a stable set of solutions differs from one robustness metric to

the next. For example, the second robustness metric stabilized very quickly, while the third and fourth robustness metrics require a substantially larger number of function evaluations. A first reason for this is that the ε values are different for non-comparative metrics. A second explanation is the fact that the first two robustness metrics use a single metric, while metrics three, four and five use two metrics. This effectively doubles the solution space.

Figure 10.3 shows the values for the decision variables for all of the solutions, grouped by metric. In this figure, each line represents a solution, and the intersection point of this line with a vertical axis shows the value of the corresponding decision variable in this solution. We observe several things. First, virtually all metrics use the edges of the solutionspace for several of the decision variables. Second, many of

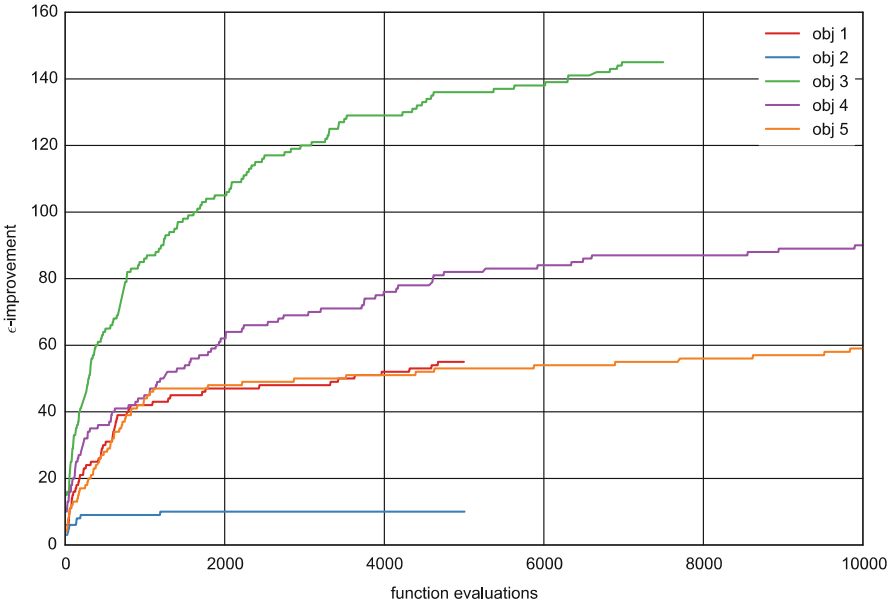


Fig. 10.2: ε-Progress for the five objective functions (robustness metrics)

the metrics go to the upper extreme of the decommissioning factor. This means that obligatory additional decommissioning of fossil-based energy generation is effective across metrics. We also see that virtually all solutions combine this with a low value for the trigger. This means that even if there are small differences between the desired and forecasted fraction of renewables, the various robustness metrics favor aggressive decommissioning. On the other levers, the pattern is less clear. Still, we observe that for example metric 1 uses the upper extreme of the desired fraction of renewables in combination with aggressive additional commissioning of sustainable energy generation technologies. Metric 1 aims at minimizing the worst case, so to minimize the worst case, very aggressive promotion of sustainable energy should be pursued. Similar observations can be made for the other metrics.

Solving the multi-objective optimization problem for each of the metrics produces a set of solutions that are non-dominated. Figure 10.3 shows the values for the decision variables for each solution per robustness metric. There are clear differences in this, but how does this translate to differences in outcomes? To analyze this, we look at the cumulative distribution of the terminal values for the outcomes of interest. This is shown in Fig. 10.4. We see that the shape of the distribution is quite similar for all solutions found by each of the five metrics, but some metrics produce better results than others. For example, if we look at the fraction of renewables (Fig. 10.4a),we see that robustness metric 5 produces a set of solutions that all

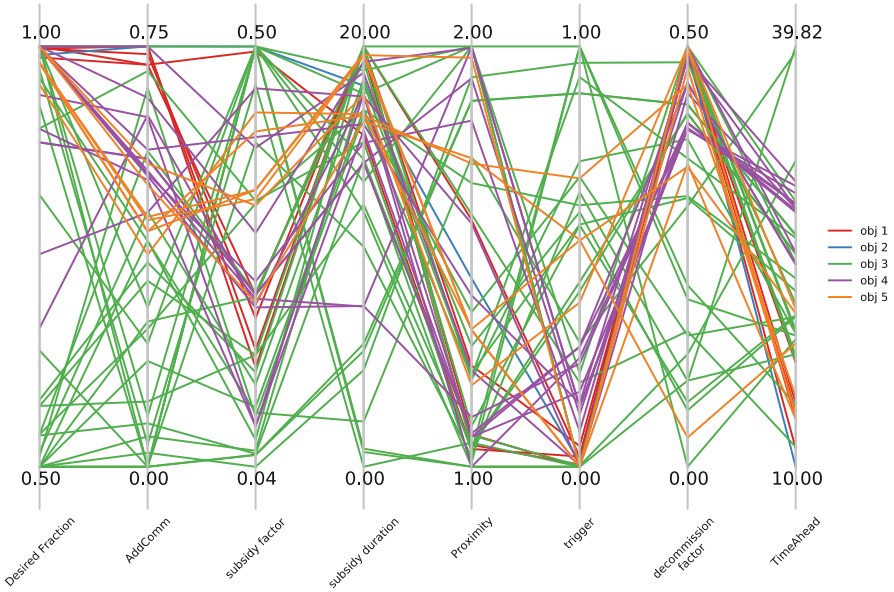


Fig. 10.3: Values for policy levers for all solutions, *colored* by objective function (robustness metric) used

result in a high fraction of cases with substantial adoption of renewables. In contrast, metrics 3 and 4 produce a set of solutions with a wider spread of results.

In order to get sharper insight into how different the results are for all of the solutions, we calculate the distance between all cumulative distributions using the Kolmogorov-Smirnov statistic. The result for each of the three outcome indicators is shown in Fig. 10.5a–c. Each of these figures shows a matrix, and each cell is colored according to the difference between the cumulative distributions of the two solutions in the row and column of this cell. Here, we can see that the one solution found by using metric number 2 (marked as obj 2–1) is equal to or better than all other solutions. The second best set of solutions is produced by metric 5 where almost all solutions are very close to the solution found by metric 2. This observation is true for all three indicators. So both metric 2 and 5 produce solutions that are

both relatively cheap, and effective in reducing CO_2 emissions and increasing the adoption of sustainable energy generation technologies.

So far, we have focused on the individual outcomes of interest. A major advantage of the multi-objective optimization approach is that we can get insight into the trade-offs between the various outcomes of interest. To this end, we created a set of pair-wise scatter plots for each of the five robustness metrics, with Gaussian kernel density estimates on the diagonal. The resulting set of figures is shown in Fig. 10.6. We observe that irrespective of the robustness metric, the solutions with a high adoption of sustainable energy generation also tend to be cheaper across virtually all scenarios. This is explained by the fact that pursuing high adoption requires aggressive

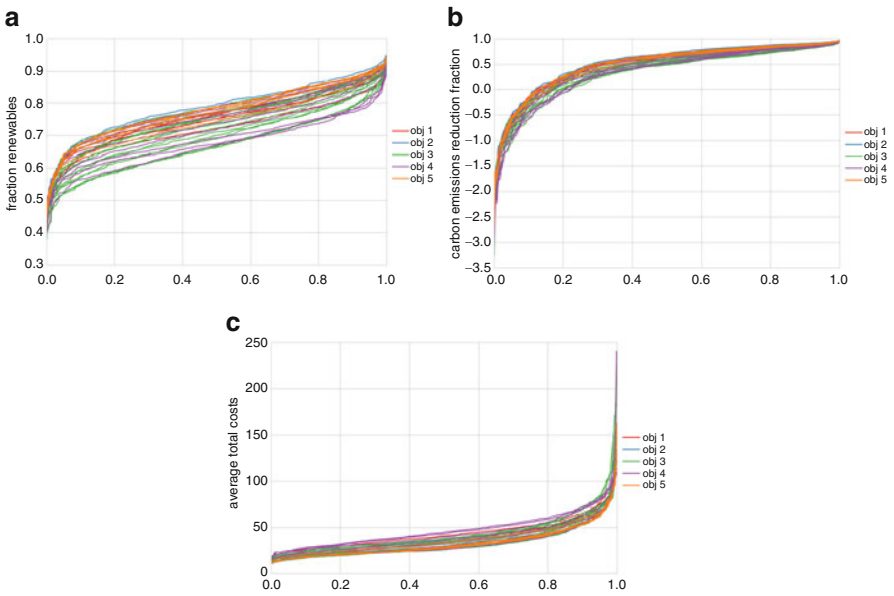


Fig. 10.4: Cumulative distributions of the terminal values of the three performance indicators across all solutions. (**a**) The fraction of renewables. (**b**) The fraction of emission reduction. (**c**) The costs of the policy

strategies in the short term which result in learning effects, in turn lowering the price and increasing the efficiency of sustainable energy generation technologies.

Figure 10.6 also highlights some of the key differences between the various metrics. For example, metric 1, which focuses on the worst case only, performs poorly across the range of scenarios even though its worst case is the least worst case possible. Metrics 3, 4, and 5 which include both the average performance as well as a metric related to the deviation from this average produce a larger set of solutions. For example, metric 4 offers some examples of solutions that appear to have a slightly lower average performance, but are more reliable with respect to their expected performance, or at least have a less wide range of downside outcomes. This is most clearly observable for the fraction of renewables.

10.5 Discussion

Comparing the results identified by the five different robustness metrics, we observe several things. First, metric 1 focuses on minimizing the worst case, but at the expense of the performance in all other scenarios. A worst case approach should thus be used only in very specific circumstances. Second, metrics 2 and 5, both of which focus on the peakedness of the distribution appear to produce the best sets of solutions. For this case, there appears to be little benefit to using a metric that focuses on the downside risk of outcomes instead of the overall distribution of outcomes. Still, both metric 3 and 4, which do this, can produce results on par with metrics 2 and 5. We see in Fig. 10.5 that both solution 4–11 and 3–16 are on par with the best solutions produced by metrics 2 and 5.

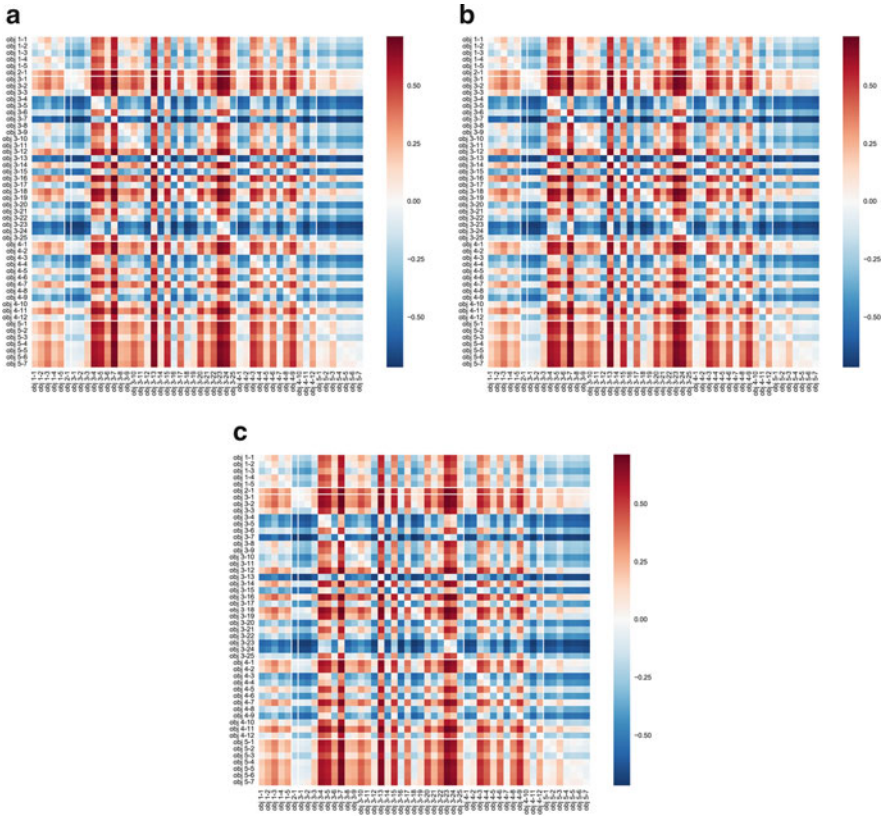


Fig. 10.5: Distances between cumulative distributions of the terminal values for the three outcome indicators as measured by the Kolmogorov-Smirnov statistic. (**a**) The fraction of renewables. (**b**) The fraction of emission reduction. (**c**) The costs of the policy

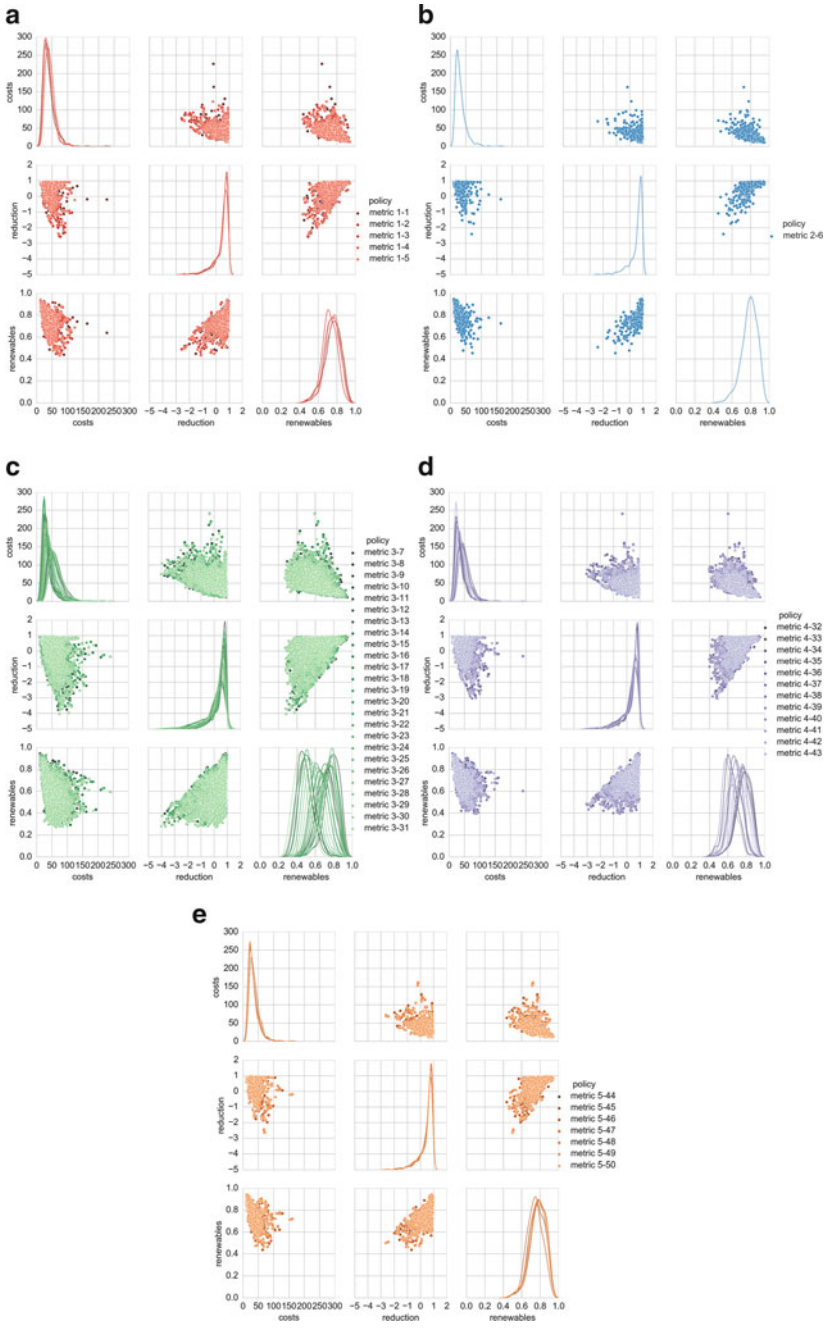


Fig. 10.6: Pairwise scatter plots for each solution across all scenarios for all five robustness metrics. (**a**) Metric 1, (**b**) metric 2, (**c**) metric 3, (**d**) metric 4, (**e**) metric 5

The results suggest that there is no single best robustness metric. Metric 1 is useful if the worst case is the sole concern. Metric 2 and 5 are useful if one is interested in reducing the uncertainty about expected outcomes but one is indifferent with respect to upside and downside deviations. Metrics 3 and 4 are useful if one is primarily concerned about negative deviations from the average performance. Metrics 3, 4, and 5 produce insight into the trade-offs between average performance and the deviations from this average. This can be very useful decision support information, but it doubles the solution space. In the present chapter, we have used the same robustness metric for all outcome indicators. Evidently a mixed approach can be pursued. For example, users might prefer certainty about costs, while being concerned about negative deviations for adoption of sustainable energy generation. In this case either metric 2 or 5 can be used for costs, while either metric 2 or metric 3 would be used for the fraction of renewables. In short, case specific considerations and system characteristics should be considered in choosing statistical robustness metrics.

The results are based on a single run of the optimization algorithm for each metric. It is good practice to assess the adequacy of solutions found through genetic algorithms by performing several replications. Genetic algorithms exploit stochasticity for effective searching, but this also introduces some randomness in the algorithm. Performing several replications with different random seeds can enhance the confidence that the identified solutions are indeed good approximations of the true Pareto front.

Robust optimization requires evaluating the performance of a given solution over a set of scenarios. This creates substantial runtime concerns. In the case reported here, we evaluated the robustness over a set of 500 scenarios. Each of these scenarios requires the running of the simulation model, which takes a few seconds. Finding effective ways of reducing the size of the set of scenarios needed for calculating the robustness metric can help in substantially reducing the calculation time. If we are able to reduce the size of the set from a 500 to 250, the runtime would be halved.

In this chapter, we focused on five examples of statistical robustness metrics. We have not considered satisficing or regret based metrics. It will be quite interesting to extent the presented analysis to also include examples of both families of robustness metrics. This would offer a more comprehensive insight into the merits of examples of all three families of robustness metrics.

The implication of the results presented in this chapter is twofold. First, there is no clearly superior single robustness metric. Case specific consideration and system characteristics affect the merits of the various robustness measures. This implies that an analyst has to choose carefully which robustness measure is being used and assess its appropriateness. Second, because of the different insights generated by the different robustness metrics, it is advised to consider multiple robustness measures simultaneously and explore their joint implications for decision-making. For example, in this case, all metrics favor aggressive additional decommissioning of existing fossil fuel based energy generation. This might be a strong argument for including this action as part of the overall strategy. In short, when deciding on robustness metrics, use multiple and choose with care.

References

1. Bai, D., Carpenter, T., Mulvey, J.: Making a case for robust optimization models. Manag. Sci. **43**(7), 895–907 (1997)
2. Ben-Haim, Y.: Information-Gap Decision Theory: Decisions Under Severe Uncertainty. Academic, London (2001)
3. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. Math. Oper. Res. **23**(4), 769–805 (1998)
4. Ben-Tal, A., Nemirovski, A.: Robust solutions of linear programming problems contaminated with uncertain data. Math. Program. **88**(3), 411–424 (2000)
5. Bertsimas, D., Sim, M.: The price of robustness. Oper. Res. **52**(1), 35–53 (2004)
6. Bryant, B.P., Lempert, R.J.: Thinking inside the box: a participatory, computer-assisted approach to scenario discovery. Technol. Forecast. Soc. Chang. **77**(1), 34–49 (2010)
7. Commission, E.: Europe 2020 a strategy for smart, sustainable and inclusive growth. COM **2020**(33) (2010)
8. Davidson, R.: Statistical inference in the presence of heavy tails. Econ. J. **15**(1), C31–C53 (2012)
9. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**(2), 182–197 (2002)
10. Dessai, S., Hulme, M., Lempert, R., Pielke, R.: Do we need better predictions to adapt to a changing climate? EOS **90**(13), 111–112 (2009)
11. Forrester, J.W.: Industrial Dynamics. MIT, Cambridge (1961)
12. Hadka, D., Reed, P.: Borg: An auto-adaptive many-objective evolutionary computing framework. Evol. Comput. **21**(2), 231–259 (2013)
13. Hamarat, C., Kwakkel, J.H., Pruyt, E.: Adaptive robust design under deep uncertainty. Technol. Forecast. Soc. Chang. **80**(3), 408–418 (2013)
14. Hamarat, C., Kwakkel, J.H., Pruyt, E., Loonen, E.T.: An exploratory approach for adaptive policymaking by using multi-objective robust optimization. Simul. Model. Pract. Theory **46**, 25–39 (2014)
15. Herman, J.D., Reed, P.M., Zeff, H.B., Characklis, G.W.: How should robustness be defined for water systems planning under change? J. Water Resour. Plan. Manag. **141**(10), 04015012 (2015)
16. Joanes, D.N., Gill, C.A.: Comparing measures of sample skewness and kurtosis. J. R. Stat. Soc. D (The Statistician) **47**(1), 183–189 (1998)
17. Kasprzyk, J.R., Nataraj, S., Reed, P.M., Lempert, R.J.: Many objective robust decision making for complex environmental systems undergoing change. Environ. Model. Softw. **42**, 55–71 (2013)
18. Kollat, J.B., Reed, P.M.: Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. Adv. Water Resour. **29**(6), 792–807 (2006)
19. Kouvelis, P., Yu, G.: Robust Discrete Optimization and Its Applications. Kluwer Academic Publishers, Dordrecht (1997)
20. Kwakkel, J.H., Pruyt, E.: Exploratory modeling and analysis, an approach for model-based foresight under deep uncertainty. Technol. Forecast. Soc. Chang. **80**(3), 419–431 (2013)
21. Kwakkel, J.H., Haasnoot, M., Walker, W.E.: Developing dynamic adaptive policy pathways: a computer-assisted approach for developing adaptive strategies for a deeply uncertain world. Clim. Chang. **132**, 373–386 (2015)
22. Lempert, R.J., Collins, M.T.: Managing the risk of uncertain threshold responses: comparison of robust, optimum, and precautionary approaches. Risk Anal. **27**(4), 1009–1026 (2007)
23. Lempert, R.J., Popper, S., Bankes, S.: Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis. RAND Corporation, Santa Monica (2003)
24. Loonen, E.: Exploring carbon pathways in the EU power sector, using exploratory system dynamics modelling and analysis to assess energy policy regimes under deep uncertainty. Master's thesis, Delft University of Technology, Delft (2012)

25. Mulvey, J.M., Vanderbei, R.J., Zenios, S.A.: Robust optimization of large-scale systems. Oper. Res. **43**(2), 264–281 (1995)
26. Popper, S.W., Griffin, J., Berrebi, C., Light, T., Min, E.Y.: Natural Gas and Israel's Energy Future: A Strategic Analysis Under Conditions of Deep Uncertainty. RAND, Santa Monica (2009)
27. Pruyt, E.: Small System Dynamics Models for Big Issues: Triple Jump Towards Real-World Complexity. TU Delft Library, Delft (2013)
28. Ray, P.A., Watkins, Jr. D.W., Vogel, R.M., Kirshen, P.H.: Performance-based evaluation of an improved robust optimization formulation. J. Water Resour. Plan. Manag. **140**(6), 04014006 (2013)
29. Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B.: Evolutionary multiobjective optimization in water resources: the past, present, and future. Adv. Water Resour. **51**, 438–456 (2013)
30. Savage, L.: The Foundations of Statistics. Wiley, New York (1954)
31. Schneller, G.O., Sphicas, G.P.: Decision making under uncertainty: Starr's domain criterion. Theor. Decis. **15**(4), 321–336 (1983)
32. Starr, M.K., Starr, M.K.: Product Design and Decision Theory. Prentice-Hall, Englewood Cliffs (1963)
33. Sterman, J.D.: Business Dynamics–Systems Thinking and Modeling for a Complex World. McGraw-Hill, Irwin (2000)
34. Takriti, S., Ahmed, S.: On robust optimization of two-stage systems. Math. Program. **99**(1), 109–126 (2004)
35. Voudouris, V., Matsumoto, K., Sedgwick, J., Rigby, R., Stasinopoulos, D., Jefferson, M.: Exploring the production of natural gas through the lenses of the ACEGES model. Energy Policy **64**, 124–133 (2014)
36. Walker, W.E., Haasnoot, M., Kwakkel, J.H.: Adapt or perish: a review of planning approaches for adaptation under deep uncertainty. Sustainability **5**(3), 955–979 (2013)

Chapter 11
Developing Robust Climate Policies: A Fuzzy Cognitive Map Approach

Alexandros Nikas and Haris Doukas

Abstract Climate change has been considered one of the most significant risks for sustainability in our century; in order to move towards low-carbon and climate resilient economies, fundamental changes must take place. In this direction, the European Union has set ambitious goals regarding the transition of its Member States to low carbon societies, but the policy strategies to promote this transition must be socially acceptable and supported. So far, climate policies have been evaluated using quantitative methods, including general equilibrium and integrated assessment models but, despite their undoubted contribution to climate modeling, both the quantitative frameworks used for studying climate change and its impacts and those aiming at policy optimization or evaluation feature significant uncertainties and limitations. In order to overcome these issues, a Fuzzy Cognitive Map based approach is proposed, aiming to directly involve stakeholders and assess human knowledge and expertise. The suggested methodological framework can significantly support climate policy making, by supplementing quantitative models and exploring impacts of selected sets of policies, based on qualitative information deriving from a structured stakeholder engagement process. Finally, an innovative approach of incorporating the concept of time into the methodology is proposed and evaluated, in the aim of enhancing the robustness of transition pathways.

A. Nikas (✉) • H. Doukas
School of Electrical and Computer Engineering, National Technical University of Athens, Iroon Politechniou 9, 157 80 Athens, Greece
e-mail: anikas@epu.ntua.gr; h_ doukas@epu.ntua.gr

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_11

11.1 Introduction

Climate change has long been considered one of the most prominent sustainability problems [70] as well as among the most significant systematic risks for global society [24] of this century. In this respect, the European Union has set ambitious goals with regard to the necessary transition of Member States to low carbon economies. The European Council has, in fact, recently mapped its envisioned requirements of a 80–95 % cut on emissions by 2050 [17], taking into account that in order to tackle the global challenge for building a sustainable future that is climate change [34] the global temperature increase should be controlled: until recently, a global average temperature increase of no more than 2 °C compared to pre-industrial levels until 2100 had been considered relatively [28, 61] safe in order for the most significant consequences to be avoided [67], although the pursuit of a newer and stricter goal of 1.5 °C was discussed and agreed upon during the latest United Nations conference on climate change [68].

It is obvious that the desired transitions cannot be achieved immediately, but require specific courses of action that can adapt to the reference concentration (and corresponding emission) trajectories, i.e. pathways, studied by the Intergovernmental Panel on Climate Change [63]. Furthermore, these transitions require radical and rapid implementation of policies that are socially and politically supported [76]. To this end, the European Union has been looking at cost-efficient ways to cut most of its greenhouse gas emissions by identifying the key sectors in this direction, as well as corresponding sectoral goals by the end of every decade until 2050, so as to achieve greater depth on costs, trade-offs and uncertainty when examining policy options [16].

So far, climate mitigation has been mostly studied through the use of quantitative methodologies. These include computable general equilibrium models that can evaluate impacts of policy reforms on the economy, revolving around the economic impacts of mitigation policies (e.g., [72]) or studied alongside climate-specific models (e.g., [40]); advanced econometric models with strong empirical background [4, 59]; dynamic stochastic general equilibrium models that emphasize economy dynamics over time, used as climate mitigation assessment tools [13, 14]; and energy-specific scenario frameworks, such as the LEAP framework [29, 42]. These models can greatly help explore implications of various scenarios in relation to climate mitigation pathways and identify effective policy drivers in quantitative terms; however, they may limit their scope to the economics of climate change, feature significant uncertainties, or fail to study all sectors in need of policy reform, and usually ignore the socio-economic dimensions of climate change.

In order to deal with the above weaknesses, integrated assessment models, that is models that draw on knowledge from multidisciplinary research [75], have been receiving increasing attention (e.g., [23, 46, 49]). These models combine economic and scientific aspects of climate change and have been proven to address issues such as evaluating climate change control policies, integrating multiple disciplines in the same framework and studying climate change in the context of both other environmental and non-environmental problems, but the degree to which their results have supported policy making is unclear [37].

Despite their undoubted contribution, all of the aforementioned quantitative models involve assumptions and simplifications [73] and, although their level of detail is limited by computing power and the need to avoid becoming too complex [47], they eventually tend to be technically too complex to construct or understand [1]. As a result, policy makers may view these models as too complicated to transform their findings into policies or as black boxes and be reluctant to trust their results [37].

It is therefore of vital importance that methods able to successfully model complex systems and at the same time easy to build by utilizing existing knowledge and experience be used. Such methods should allow for both feasible and successful policy pathways to be determined and studied, taking into account the national context and sectoral specificities of a country as well as incorporating the knowledge of experts around uncertainties and risks where quantitative data are not available or too costly.

In this study, Fuzzy Cognitive Maps (FCMs) are proposed as one such methodology, for modeling the complex system that is climate mitigation policy in the European Union. FCMs are fuzzy structures that strongly resemble neural networks and are often used as a useful tool for modeling complex systems [35]. Once constructed, the FCM model allows performing qualitative simulations of a system and experimenting with the model [2]. FCM simulations thus allow decision makers to examine information dynamics and uncertainty, as well as identify critical system elements and assess different decision alternatives by comparing their outcome in a holistic manner [33].

Özesmi and Özesmi [51] extensively discuss the reasons for choosing FCMs when dealing with environmental problems over other modeling methods, such as system dynamics models, multiple criteria decision analysis methodologies, expert systems and structural equation models. Most importantly though and compared to quantitative models in particular, FCMs do not depend on data availability: their robustness does not depend on any training procedures that are biased to the size of the available data sets [53]. Moreover, given the fact that they are built on human expertise and knowledge alone, they are highly flexible and easy to include social effects [69].

The following section introduces the origins and structure of Fuzzy Cognitive Maps. Section 11.3 presents the proposed FCM-based methodological framework for modeling climate policy pathways. The aim of the framework is to visualize the system dynamics deriving from both the results of other modeling methods and an effective stakeholder participatory process; quantify the knowledge and experience of the experts with regard to the success of the various transition pathways; and use all available information to simulate the impacts of the chosen policies on the respective systems, in order to gain new insight. Finally, Sect. 11.4 focuses on the assessment of simulation results, and Sect. 11.5 discusses key aspects of the proposed methodological framework.

11.2 Fuzzy Cognitive Maps

Cognitive mapping is a qualitative technique that aims to capture a person's perception of a particular issue in a diagrammatic format [20]. It provides both the analyst and the interviewee with a process that is not constrained by formal structure and through which additional insight can be gained [12]. In this context, a map focuses on the individual's values, beliefs and assumptions about a certain domain and reveals how these relate to each other, providing information about how the change of one issue can affect the others. Therefore, a cognitive map [3] can be defined as the graphical representation of a system, in which nodes represent concepts and arcs represent the perceived relationships between these concepts (Fig. 11.1).

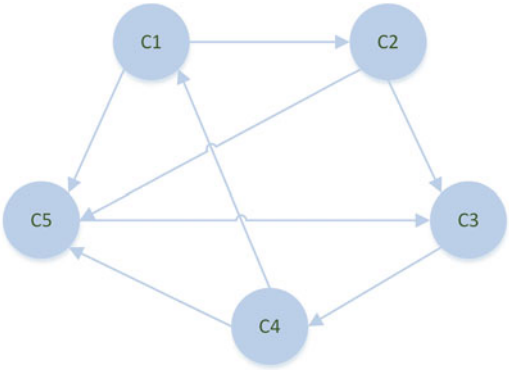


Fig. 11.1: Example of a cognitive map

Every cognitive map features a unique adjacency matrix $\mathbf{A} = [a_{ij}]$ that provides information with regard to its structure, i.e., how concepts are linked to each other [27]. The adjacency matrix is a square matrix that includes all concepts listed on both the vertical and the horizontal axis; when a causal connection from concept c_i to concept c_j exists, then $a_{ij} = 0$, otherwise $a_{ij} = 1$. In other words, if $a_{ij} = 1$, then concept c_i is considered a cause of concept c_j, and concept c_j is considered an effect of concept c_i. The adjacency matrix for the cognitive map in Fig. 11.1 is shown in Table 11.1.

Cognitive maps work as a transitional object applied by members in the aim of expressing and understanding their knowledge contents with regard to certain problem domains, as well as their structure [6]. They can be used for assessing attention, association and importance in order to identify mental connections between strategic themes; showing dimensions of categories and cognitive taxonomies in order to determine hierarchical relationships and frame the competitive environment; exploring influence, causality and system dynamics; delving into the structure of arguments and conclusion; and specifying schemas and frames [31].

Table 11.1: The adjacency matrix of the example cognitive map

	C_1	C_2	C_3	C_4	C_5
C_1	0	1	0	0	1
C_2	0	0	1	0	1
C_3	0	0	0	1	0
C_4	1	0	0	0	1
C_5	0	0	1	0	0

Kosko [39] was the first to introduce the notion of Fuzzy Cognitive Maps (FCMs), by suggesting that cognitive maps are too binding for knowledge-base building because causality is fuzzy, admits of vague degrees and thus cannot be fully described by arcs that connect concepts with one another. Among Huff's five types of cognitive maps, FCMs fall under the third category: they are fuzzy-graph structures for representing causal reasoning and consist of concepts that interact with each other, enabling the mapper to show the dynamics of a particular system [25]. The main difference between FCMs and cognitive maps lies in the fact that causal relation values are also defined and quantified, meaning that links between concepts are weighted.

As a result, a fuzzy cognitive map does not have only a unique adjacency matrix but a weight matrix $\mathbf{W} = [w_{ij}]$ as well. Entries in an FCM weight matrix are not of binary form (either 0 or 1), but can be of any numerical value within the interval $[-1,1]$. If there exists a causal connection from concept c_i to concept c_j, then $[w_{ij} \in (0,1]$ if a positive change in concept c_i leads to an increase in concept c_j, or $[w_{ij} \in [-1,0)$ if a positive change in concept c_i leads to a decrease in concept c_j; otherwise, if no connection exists between the two concepts then $w_{ij} = 0$.

A causal map, and therefore a fuzzy cognitive map as well, includes concepts that can be one of three types: transmitters, receivers and ordinary concepts [21]. A transmitter is a concept that can be perceived only as a cause to other elements within the system under examination; a receiver is a concept that can be perceived only as an effect of other elements; and ordinary concepts are those elements that have both at least one cause and at least one effect relationship. These three types can also be defined by their indegree and outdegree functions. The indegree or generalized indegree of a node is the number of paths leading to this particular node from others, while the (generalized) outdegree of a node is the number of paths leading from this node to others [9]. As a result, a transmitter is a concept with zero indegree and non-zero outdegree; a receiver is a concept with non-zero indegree and zero outdegree; and an ordinary concept is a concept with both indegree and outdegree non-zero [50]. An example of these structural criteria can be viewed in Fig. 11.2, which is inspired by the work of Lopolito et al. [45].

Other than assessing the structure of a map, the indegree and outdegree functions can also contribute to drawing comparisons between FCMs, as they are used to calculate centrality of concepts, which is the summation of their indegree and

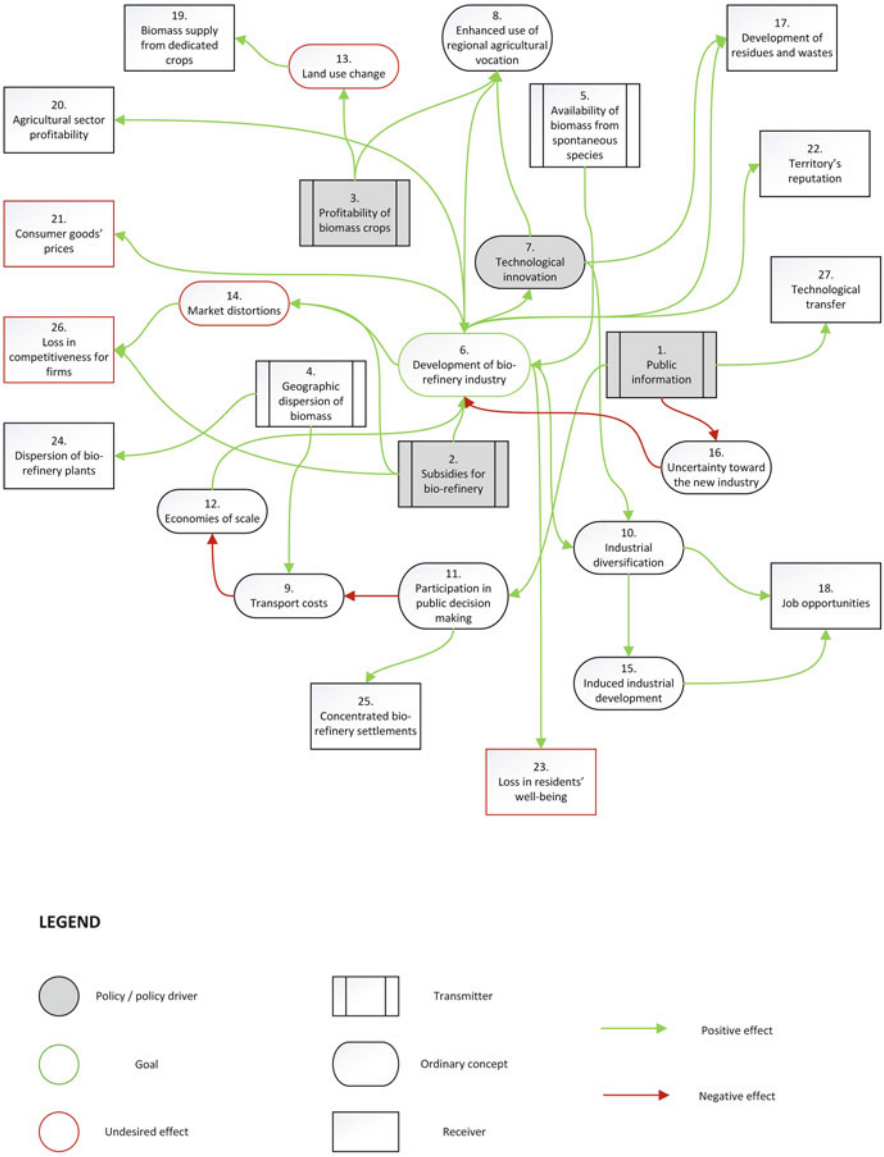


Fig. 11.2: Transmitter, receiver and ordinary variables of a causal map, based on Lopolito et al. [45]

outdegree [27], as well as complexity, which is defined as the ratio of number of receivers to the number of transmitters [51].

In a fuzzy cognitive map, concepts represent key factors and stand for events, goals, inputs, outputs, states, variables and trends of the modeled system [25]. For example, in the context of developing robust mitigation policy pathways, these can

be events, policy-defined goals, system trends, transition drivers and barriers (such as risks and uncertainties) and other variables acting within each national (or sectoral) system.

After a fuzzy cognitive map has been drawn, using techniques from neural networks, the systematic causal propagation [39] of the map can be analytically traced through a simulation process [52]. Using a simulation driver function, the value of a concept during each iteration depends on its value at the beginning of the current step and the values of the concepts that causally point to it, as well as the causal weights of the respective interconnections. At the end of each iteration, new values are normalized within the interval $[0, 1]$ using a transfer function (also known as threshold or transformation function). Simulations may converge to a steady state vector (fixed point), a limit cycle where the concept values fall into a loop of numerical values, a limit torus, or a highly unstable strange (chaotic) attractor in the fuzzy cube [19], where concept values randomly reach varying values. The result to which the simulated system will converge depends on the initial state vector, given a fixed model structure.

This analysis allows the modeler to explore what-if scenarios, by performing simulations of the FCM for various initial state vectors, i.e., for different sets of activated policies or different levels of activation of certain policies. Comparisons between the results can be used to support decision making or scenario building [62].

It is noteworthy to mention that there exist many applications of Fuzzy Cognitive Mapping in literature, with regard to climate change and environmental planning and assessment, among which many focus on improving the methodological process as well, while others aim to either stress the participatory methodology or exploit the FCM methodology for building scenarios [48].

11.3 The Methodological Framework

Regardless of how FCM simulation results have been evaluated and supported decision and policy making, FCMs have so far been used to model and simulate various systems from different disciplines, in problems that feature significant uncertainties linked to the social factor to some (e.g., [2, 45, 52, 78]), great (e.g., [8]) or no extent at all (e.g., [43, 64]). When considering the many risks and uncertainties associated with climate change policy [22] as well as issues concerning public acceptance of low-carbon technologies and policies, one can understand the logic behind using fuzzy cognitive mapping in this context. However, as already discussed and unlike many of the domains in which FCMs have previously been used, there exist a large number of quantitative frameworks that can support climate mitigation policy making and that have, to some degree, tackled many uncertainty issues.

The aim of this approach, therefore, is not to avoid the necessary quantitative models used for optimizing, evaluating and selecting policies that can deliver the desired transition goals, but rather supplement quantitative methodologies, by linking them to qualitative, experience-driven modeling. Fuzzy cognitive mapping

does not provide any real-value estimations [52]: FCM simulation is based on purely qualitative information and is therefore not intended for forming exact quantitative values [6], nor should it be perceived as such, but rather as a means of exploring which of the proposed policies or sets of policies are *believed* to perform better in the examined systems.

Using fuzzy cognitive maps for linking stakeholder input and quantitative modeling frameworks has been explored in the past [69], in the aim of enabling the quantification and integration of narrative storylines, i.e. participatory output, into the models. This approach assumes the need for the opposite sequence: after having determined a number of policy pathways that according to quantitative models results are sufficient to achieve the required transitions, FCMs can help select the optimal pathways, by utilizing expert knowledge and experience as to the feasibility and applicability of the different sets of policies. This is similar to the process Hobbs et al. [30] suggested undertaking in the Lake Erie Lakewide Management Plan project, following a fuzzy set theory approach, before considering the FCM alternative.

The proposed framework consists of the following steps, presented below:

- **Step 1.** Determining the group of stakeholders
- **Step 2.** Designing the cognitive map
- **Step 3.** Inferring causal relation weights
- **Step 4.** Exploring the time dimension
- **Step 5.** Quantifying concepts
- **Step 6.** Selecting configuration parameters
- **Step 7.** Running simulations

11.3.1 Determining the Group of Stakeholders

For the purpose of evaluating potential climate mitigation policy pathways within each country, experts from the following groups are to participate in the process:

– Government departments
– Private and public sector industries, associations and distributors
– Policy makers
– Research institutes
– Non-Governmental Organizations (NGOs) representing various societal groups, such as consumers and environmental protection activists
– Labor unions
– International organizations
– Electric utilities and regulators
– Media
– Communities and households

All country-specific stakeholders take part in the framework in the same way, responding to common surveys, questionnaires and workshops; however, they are nevertheless grouped into respective categories, in order to better assess the qualitative information they contribute.

11.3.2 Designing the Cognitive Map

This second step of the framework is the most challenging one. Normally, this step starts from scratch with experts being asked to either help the analysts design the map or draw the map themselves. Özesmi and Özesmi [51] discuss most of the existing ways in which cognitive maps can be obtained.

The most common way of doing that, in FCM literature, has been the use of questionnaire-oriented interviews or meetings, through which experts express their perceptions of concepts that are important with regard to the problem domain under examination. Roberts [57] explains that a list of variables is initially created and then refined, by selecting the most important ones to be included in the cognitive map. Isak et al. [32] present an analytical framework in which interviews can be conducted. Instead of separate interviews, collaborative determination of key concepts can be conducted, by organizing workshops [41].

Another commonly found process for visualizing expert input into maps has been that of experts directly drawing their cognitive maps. Özesmi [50], after enabling interviewees to express their views on the important variables of the system, had them draw lines with arrows bearing positive or negative signs between variables, showing causal increase or decrease respectively. A detailed description of an indicative framework for achieving this can be found in Özesmi and Özesmi[51], where stakeholders were guided along the whole procedure by using engaging questions. Papageorgiou and Kontogianni [52] too, after explaining the process of constructing fuzzy cognitive maps, asked each of their interviewees to draw their map.

Other methods include coding cognitive maps from implicit or explicit descriptions found in experts' narratives [77] and extracting cognitive maps from existing numerical measurements [60].

The proposed approach assumes that key concepts have been a priori determined, which however do not capture the whole picture nor include several uncertainties and insights that only experts are able to provide. Policies and transition drivers are identified in extensive case studies of the national systems of innovation [5, 58, 71] with regard to one or more sectors of interest and for each of the examined countries, while other key concepts can be extracted from the results of quantitative frameworks used for evaluating these policies onto the examined national and sectoral systems, as previously discussed. To complete the initial set of concepts, the most important risks and uncertainties, which are identified after a detailed literature review and prioritized using Multiple Criteria Decision Analysis methods, are added.

Given the large number of systems to be modeled, we select the first option presented: after determining the sets of policies to be examined, i.e. those that are able

to achieve the transitions to low carbon societies as per models results, as well as other key concepts, stakeholders are asked to provide their input during interviews. The research question that needs to be addressed at this stage is "*what does the informant perceive as important concepts in the problem domain, given the initial set of key concepts?*". By following this approach, stakeholders' opinions are taken into account in order to facilitate the development of the FCMs, while also minimizing their bias with regard to their suggestions [36]. Their liberty to add concepts is not limited but encouraged, as long as they include the predefined key concepts or, even better, are driven by the latter when listing their perceived concepts of significance. In countries where such an option exists, time-effective workshops are preferred.

After putting together a large list of concepts for a particular system, it is imperative that only a global and concise map be created. There exist various ways of refining concepts, by reducing their number to those that appear to be most critical. Lopolito et al. [45] pursued consensus over a compromise between the required preciseness and the unavoidable vagueness that the human language features as well as over the selection of the final set of concepts, during their discussion with stakeholders. Other scholars have taken advantage of structured communication methodologies in order to achieve consensus, such as the Delphi method [7]. Given the large number of different systems to be examined, our aim is to avoid a long, iterative process for each country and limit the stages of stakeholder engagement for the purpose of creating cognitive maps. To this end, the task of reducing the number of concepts is entirely up to the analysts. At first, all identified concepts are reviewed in case groups of two or more should be semantically merged and, then, the remaining concepts are prioritized by number of occurrences among stakeholders, so that concepts (including merged ones and their cumulative number of mentions) that rank at the bottom may be omitted.

Afterwards, a second round of stakeholder engagement revolves around the research question "*which of the other concepts does each concept affect, either positively or negatively?*". Experts only need to specify which of the other concepts each concept has a causal relationship with; a structured way to achieve this is to enumerate all concepts and have experts note the numbers of variables of which each concept can be considered a cause, or ideally provide them with an empty adjacency matrix and have them fill it in.

After having obtained the adjacency matrix of this refined set of concepts, more techniques are available to the analysts for obtaining a final and concise cognitive map. A popular technique in the literature has been that of replacing subgraphs with a single concept, which is called condensation [27] or aggregation [51]. For example, Papageorgiou and Kontogianni [52] describe how they managed to cluster 52 variables into 26 high-level concepts. It should be noted that the goal is not to alter the derived model but rather simplify it and optimize its complexity; according to Özesmi and Özesmi [51], when the number of variables exceed the number 30 the map starts being counterproductive for gaining insights.

The whole process of designing the cognitive map can be viewed in Fig. 11.3.

At this point, it should be noted that, as far as receiver and ordinary variables are concerned, only concepts that feature a cause and/or effect relationship with the

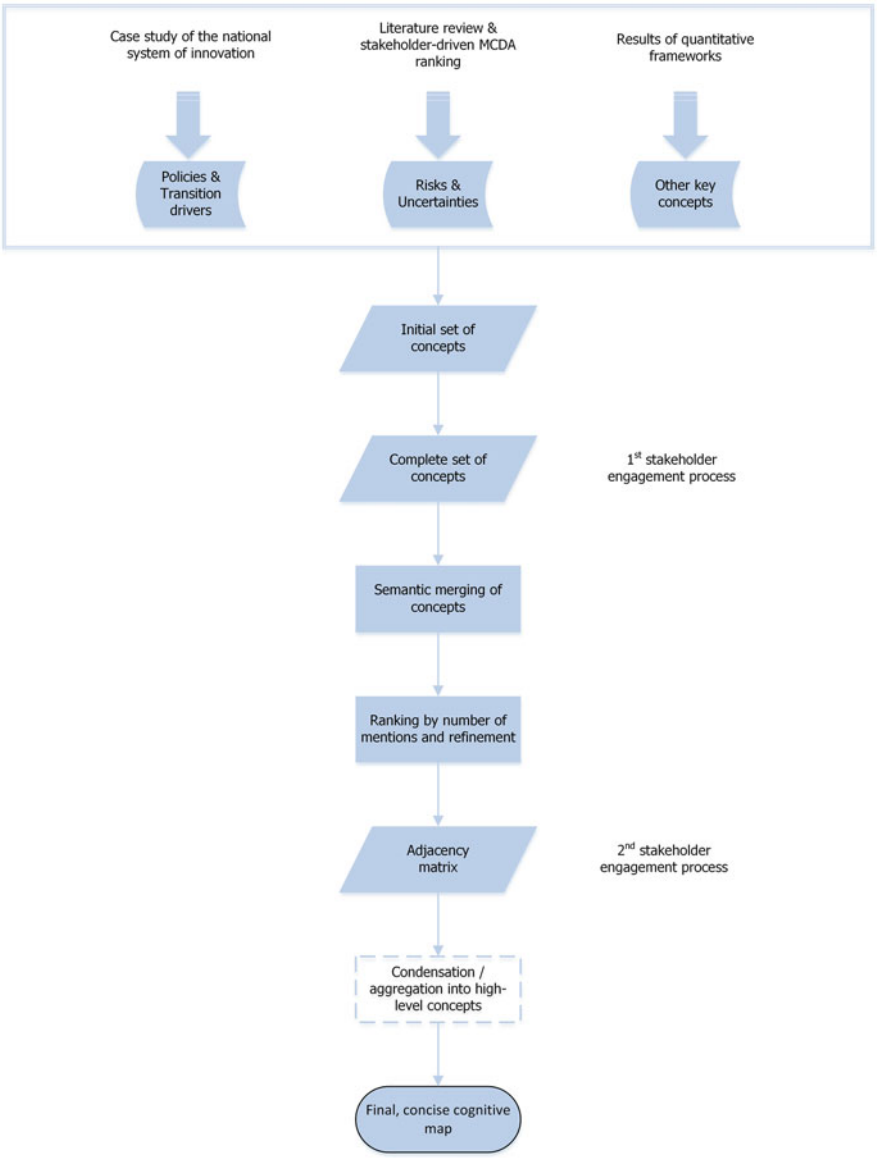


Fig. 11.3: Flow chart of Step 2 of the framework: designing the cognitive map

respective system should be included in the final map, while transmitter variables should exclusively be policies, as it is the policies' effect on the system that we aim to explore. For example, although "population growth" is a concept that may have direct impact on a number of variables within a system, it should not be included in the map as the objective is to explore how the system reacts to certain policies every

time and draw comparisons, while "population growth" exists in every possible set of policies; however, if any concept of the system affects "population growth" in any way, then it should be included as well.

11.3.3 Inferring Causal Relation Weights

This stage calls for the third and final stakeholder engagement process which revolves around the research question "*given the cognitive map, as well as the national context, how strongly does each concept affect the others?*" and the purpose of which is to establish the grade of causality between concepts. In literature, these weights can be of binary form $\{0, 1\}$ depending on whether two concepts are directly related or not (e.g., [1]), take any value within the interval $[0, 1]$ (e.g., [7]), or take any value within the interval $[-1, 1]$ (e.g., [8]), which is also the case in our proposed framework. The sign expresses whether the causal concept has a positive or negative impact on the effect concept, and the value shows how much impact the former has on the latter.

There exist many different approaches for extracting the required information from experts. Özesmi and Özesmi [51], after having the stakeholders draw their own cognitive maps, asked their interviewees to directly assign numerical weights between -1 and $+1$ to the perceived interconnections. Papageorgiou and Kontogianni [52] proposed using if-then rules that infer a fuzzy linguistic variable from a given set of thirteen variables; alternatively, experts could directly assign fuzzy linguistic weights from the same set. In the same context, Lopolito et al. [45] used only three linguistic values for describing the grade of causality for the concept interconnections (namely weak, moderate and strong) and assigned these linguistic values to numerical weights: 0.33, 0.67 and 1, respectively. Other approaches include using structured communication techniques: for example, Biloslavo and Dolinšek [6] used an iterative approach with feedback based on the Delphi method [18] and the Multiple Criteria Decision Analysis method called Analytical Hierarchy Process [66] for calculating weights; Olazabal [48] used the Q-methodology [11] in order to support the whole stakeholder engagement process, including inferring weights. Last but not least, other researchers proposed iterative, penalization-based stakeholder engagement algorithms using a credibility factor [25].

Given the need to limit stakeholder engagement to the least possible number of stages, as dictated by the very large number of systems to be examined as well as the fact that this methodological framework presupposes the use of other models beforehand (Fig. 11.3), we propose that every stakeholder be free to choose among the 13 linguistic values of the following set [52]:

{negatively very very strong, negatively very strong, negatively strong, negatively medium, negatively weak, negatively very weak, zero, positively very weak, positively weak, positively medium, positively strong, positively very strong, positively very very strong}

These values, however, are directly assigned to the following numerical set:

$$\{-1, -0.83, -0.67, -0.5, -0.33, -0.17, 0, +0.17, +0.33, +0.5, +0.67, +0.83, +1\}$$

After individual weights have been inferred, all numerical values are aggregated using either the mean or the weighted mean of all weights for each causal relation; this depends on whether the analyst chooses to assign different weights to different stakeholders, as identified and grouped in Step 1 of the framework and based on the level and nature of expertise each stakeholder group has. For example, members of the media, consumers or community representatives are not as knowledgeable about every identified concept or policy as experts from research institutes or policy makers. For our purpose and given an appropriate sampling of experts, a mean value is considered adequate for aggregating causal relation weights. Finally, the weight matrix is filled in using the aggregated weight values.

11.3.4 Exploring the Time Dimension

A very weak point of Fuzzy Cognitive Mapping, as applied so far, has been that of appropriately defining and incorporating time [69]. The output of an FCM simulation can be used to assess how key factors play out after a number of iterations, when certain policies or sets of policies are at play, but this output cannot be directly translated into time. However, this can be overcome to some extent, by only including causal relations that are meaningful in the same time scale [38].

Hagiwara [26] was the first researcher to acknowledge the lack of the concept of time in FCMs, along with the need for conditional and non-linear causal relationships; he therefore proposed an extended FCM framework that is able to incorporate time delays, non-linear weights and conditional relations. However, this approach not only requires an extensive, time-consuming stakeholder engagement framework that assumes experts' knowledge and expertise is adequate for offering insight into all of the required information, but also significantly enhances the complexity of the FCM methodology.

In order to incorporate time relationships into Fuzzy Cognitive Mapping, Park and Kim [56] suggested collecting information on the time delay that every relationship between concepts features and then translating the original FCM into a fuzzy time cognitive map (FTCM) in which all causal relations apply in the same time unit. In order to achieve this, they proposed a framework that introduces dummy nodes between concepts and calculates the sign and value of individual weights that compose each original causal weight. Although this approach serves our purpose, in larger FCMs (such as the ones we anticipate to have to deal with, with regard to climate mitigation policies) there is significant increase in the FCM complexity and the visual outcome is hard to produce, supervise and gain insight from. This is partly why the proposed FTCMs only incorporate time lags of two or three time units and no more, but this poses another limitation since causal relations may potentially feature larger delays.

Other approaches include the expression of the implicit time delay of every relation and the selection of a base time in Rule-Based Fuzzy Cognitive Maps [15]; the use of Fuzzy Time Cognitive Maps for analyzing trust dynamics in virtual enterprises [74]; the agent-based FCM methodological framework developed by Lee et al. [44], in an effort to better address the drawbacks identified by Hagiwara [26] and further analyzed by Schneider et al. [60], which was applied in industrial marketing planning; and a significantly more complex version of timed fuzzy cognitive maps that requires the determination of linguistically-expressed time-dependent weights [10].

In our approach, all causal relations are assumed to be yearly and stakeholders are asked to evaluate how many time units (i.e., years) it takes for each cause concept to have an impact on its effect concept(s). The research question that needs to be addressed at this stage is "*how long does it take for a change in each concept to have an effect on its adjacent concepts?*". This process does not necessarily require an extra stakeholder engagement step, as it can be done in parallel with the weight inference stage (Step 3). After extracting this information, again an average value is calculated for each causal relation, as in Step 3.

Time lags are incorporated in the way new concept values are calculated, as described in Step 6. As discussed in Sect. 11.4, this approach does not provide different numerical results from what the FCM simulations would produce without integrating time lags, but offers new insight with regard to how soon each policy or set of policies can produce the calculated results.

This step concludes the stakeholder engagement process, which can be viewed in Fig. 11.4.

11.3.5 Quantifying Concepts

As explained in Sect. 11.2, all concept nodes are assigned a numerical value, which is usually within the interval $[0, 1]$ or $[-1, 1]$, depending on which threshold function is selected (Step 6). Certain researchers have proposed different approaches; for example, when suggesting an FCM-based approach for robust scenario planning with regard to wind energy development, Amer et al. [2] used concept values of binary form, i.e. $\{0, 1\}$, depending on whether a concept is clamped or not. Although such an approach could be applied for developing scenarios emerging from policies or sets of policies that greatly differ from one another, it would not allow for adequate comparisons in cases where policies have different effects on the same variables or where policies or sets of policies are clamped to a different extent among alternatives. In our approach, concepts are able to take values within the interval $[-1, 1]$, so as not to limit our methodology to a specific threshold function.

A major issue that has never really been addressed or elucidated in the FCM literature is that of clarifying which of the concepts should be modeled as constants and which ones as variables. Since a Fuzzy Cognitive Map model is simulated over and over again using a particular threshold function and until it converges to a steady

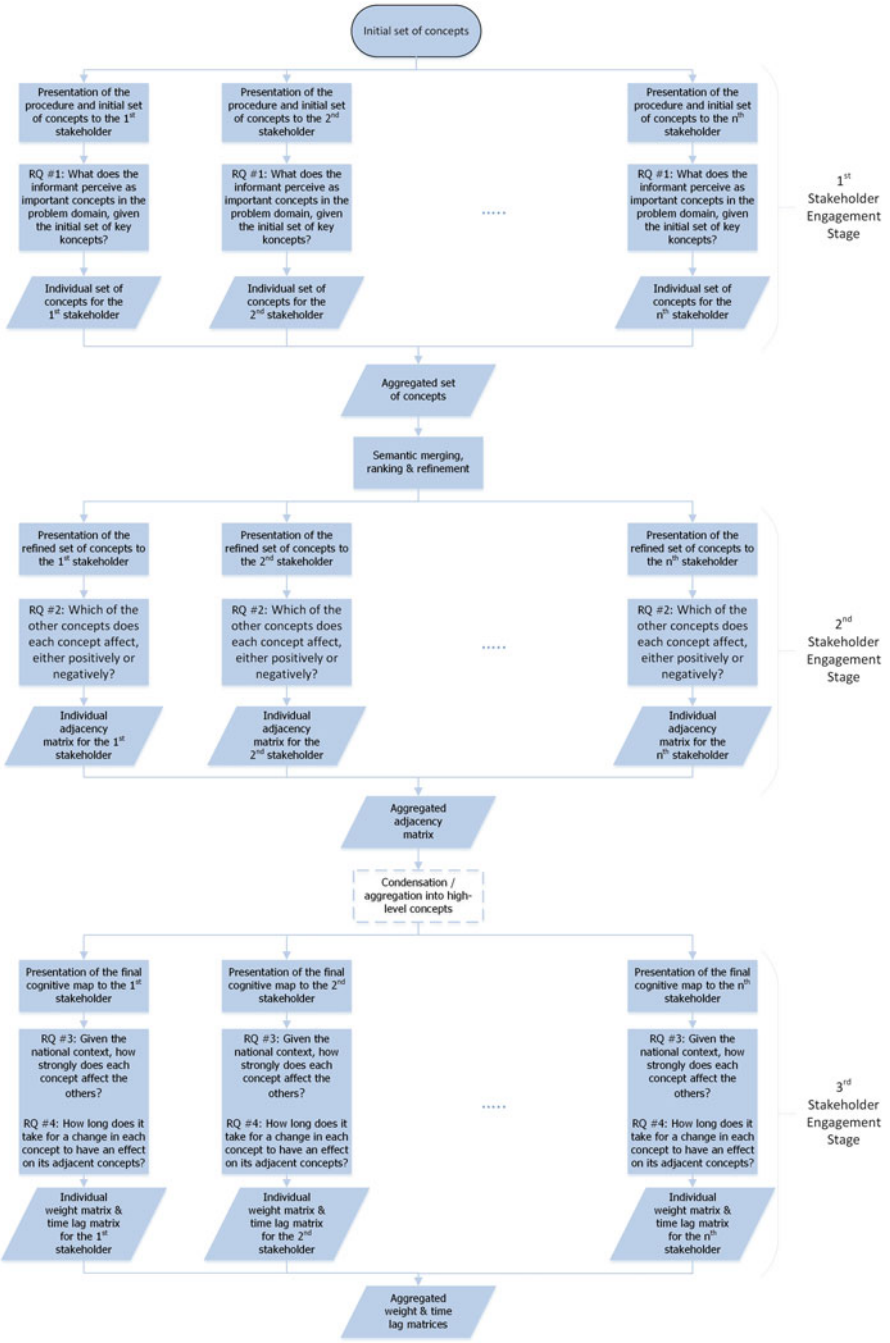


Fig. 11.4: The three stages of the proposed stakeholder engagement process

state vector, if all concepts were modeled as variables this final state vector would depend entirely on the FCM structure: set of concepts, adjacency matrix and weight matrix. In order for each simulation to depend on the initial state vector and thus be meaningful (given a fixed set of concepts, interconnections and weights), concepts that differentiate each alternative should be unaffected by the selected threshold function. As a result, policy (transmitter) concepts should be modeled as constants, while all other (ordinary and receiver) concepts should be modeled as variables.

Therefore, the initial state vector should be carefully selected for each alternative option considered and modeled. In most cases, different policies are examined and activated (i.e. 0 if unclamped and 1 if clamped) every time but this should not necessarily be the case; for example, Papageorgiou and Kontogianni [52] examined two different scenarios for future prospects and risks of the Black Sea marine environment, by attributing different, non-zero initial values to all concepts in each scenario.

11.3.6 Selecting Configuration Parameters

One simulation driver function that has almost exclusively been used in FCM applications calculates the value of a concept at the end of an iteration as the sum of contributions of its causal concepts at the beginning of the iteration:

$$C_j^{(t)} = f\left(\sum_{i=1}^{n} C_i^{(t-1)} w_{ij} + C_j^{(t-1)}\right)$$

where $C_j^{(t)}$ is the value of concept j at the end of the iteration, $C_j^{(t-1)}$ is the value of concept j at the beginning of the iteration, $C_i^{(t-1)}$ is the value of concept i at the beginning of the iteration , and f is a threshold function.

Depending on the notion of auto-correlation that we will be using, the second term ($C_j^{(t-1)}$) can be omitted; functions of this form assuming no auto-correlation have been used in FCM literature. Of course, both functions can be considered identical, depending on whether the weight matrix includes auto-correlation: by assigning ones on the main diagonal of the weight matrix, i.e. $w_{ij} = 1$, then auto-correlation is implied and included in the first term and the second term should be omitted.

Stylios and Groumpos [65] proposed a slightly modified approach for calculating new concept values at the end of every iteration, by introducing two coefficients k_1 and k_2, depending on how strongly auto-correlation affected each concept compared to the contribution of its causal concepts:

$$C_j^{(t)} = f\left(k_1^j \sum_{i=1}^{n} C_i^{(t-1)} w_{ij} + k_2^j C_j^{(t-1)}\right)$$

However, this approach too requires significantly more information from the stakeholder engagement process, since stakeholders should not only know but also be able to offer insight into the extent to which each and every single concept depends on its previous value compared to the contributions of its causes.

Another approach was introduced by Papageorgiou and Kontogianni [52], according to which limitations presented by the sigmoid threshold function, when initial values are 0 or 0.5 or the initial state cannot be efficiently described [54], can be addressed:

$$C_j^{(t)} = f\left(\sum_{i=1}^{n}\left(2C_i^{(t-1)} - 1\right)w_{ij} + 2C_j^{(t-1)} - 1\right)$$

Finally, Papaioannou et al. [55] suggested an interesting simulation driver function that, given a preliminary process of adjusting and greatly reducing weights depending on the number of iterations at which the system is expected to converge, requires no transformation function:

$$C_j^{(t)} = \sum_{i=1}^{n}\left(C_i^{(t-1)} - C_i^{(t-2)}\right)\frac{C_j^{(t-1)}}{C_i^{(t-2)}}w_{ij} + C_j^{(t-1)}$$

In the proposed framework, however, we will be using a different simulation driver function, based on the originally presented and most frequently used one, so as to incorporate the notion of time delays. This is done by multiplying the weight of a relation between causal node i and effect node j with the value of concept i at time t minus the time delay lag_{ij} of the respective impact:

$$C_j^{(t)} = f\left(\sum_{i=1}^{n} C_i^{(t-lag_{ij})}w_{ij} + C_j^{(t-1)}\right)$$

Finally, a threshold function must be selected and applied, in order to normalize new concept values within the interval $[-1,1]$ at the end of every iteration. As in neural networks, there exist many different threshold functions, such as logistic, linear threshold or step functions [51]. However, the most frequently used threshold functions in the FCM literature are the sigmoid function and the hyperbolic tangent function.

When concept values can only be positive, i.e. belong to the interval $[0,1]$, the unipolar sigmoid function ($\lambda > 0$) is usually used as a threshold function (e.g., [45]):

$$f(x) = \frac{1}{1+e^{-\lambda x}}$$

If concept values can be negative as well, meaning that they can take values within the interval $[-1,1]$, the hyperbolic tangent function can be used instead [25]:

$$f(x) = \tanh(x)$$

Having quantified concepts in the interval $[-1, 1]$, both threshold functions can be applied to our approach. The non-negative transformation of the sigmoid function allows for better understanding of concept activation levels [51]. Furthermore, given that it squashes values into a stricter interval, it usually converges faster than the hyperbolic tangent function. However, the latter allows for more realistic representation of the causal output and requires no calibration. For example, if the value of a certain concept can only decrease, the sigmoid function will still produce a positive outcome in the final state vector, which must then be compared to the inaction model in order to draw conclusions for a single scenario; of course, when dealing with a number of alternatives, comparisons between alternatives can still be drawn without comparing them to the inaction scenario, using the sigmoid function.

11.3.7 Running Simulations

After the FCM model has been developed and all configuration parameters have been selected, simulations can begin for each country case study and respective sector; for every simulation and as discussed in Sect. 11.2, causal propagation takes place in every iteration until the system converges; this happens when no change occurs in any of the concept values after a certain point, the concept vector at which is called the final state vector.

11.4 Assessing Results

After having simulated the derived model for every possible policy strategy, comparisons between the final state vectors of the alternatives should be drawn in order to assess to what extent the desired transition has been promoted by activating each set of policies. The larger the value of the goal concept is at the end of the simulation, the better the selected policies are considered by the stakeholders. Of course, the analysts must define which of the concepts included in the FCM should be taken into account when comparing the various alternatives, as well as how important each of these concepts are. For example, Lopolito et al. [45] defined the concept representing the policy objective and distinguished five other concepts to negatively contribute to the alternative pathways' score, as equally undesired effects of the policies. Özesmi [50] proposed a different approach: different objectives were acknowledged for each of the stakeholder groups, leading to different conclusions for each group's perspective. On the other hand, when comparing scenarios with strict activation thresholds and values of binary form, all of the final state vector values are taken into account in order to produce sound narratives (e.g., [2]).

Other than the final state vector and given the integration of the time dimension into the proposed framework, another significant criterion to be taken into consideration when comparing the examined alternatives is how fast convergence has occurred.

As an example, simulations were run for the fuzzy cognitive map presented in Fig. 11.2, with and without time delays, and using the original weight matrix. When activating the first policy option, i.e. enhancing public information on the bio-refinery industry, all concept changes (including the desired transition) appear to slightly delay when very little time lag is introduced in certain interconnections, compared to a no-lag model (Fig. 11.5). It should be pointed out that the sigmoid threshold function was used.

In order to achieve a better understanding and more detailed examination of the time lags' effects, the same model was simulated using the hyperbolic tangent threshold function, using both of the options previously considered and adding a third one with a slightly increased time lag. The impact of both time delays on the goal concept, which is the development of the bio-refinery industry, can be viewed in Fig. 11.6.

Looking at both figures, it is safe to conclude that time delays do not alter the values of the final state vector, as produced without integrating time lags in the first place, but nevertheless offer new insight into how fast changes take place. As previously discussed, the final state vector depends solely on the FCM structure, including its concepts, interconnections and weights, as well as the initial values of the considered transmitter (policy) concepts. As a result, time delays introduced in the proposed manner cannot change the results, unless a certain delay is perceived to be so prolonged that the respective causal relation's impact is not observed before the system converges; this is a negative feature of the proposed approach that could potentially lead to false results, but which can easily be addressed by either comparing results with the respective no-lag model's results or introducing stricter convergence criteria, such as requiring zero numerical changes in the state vector for a number of iterations equal to the largest time lag included.

As expected, the observed delay had an impact on convergence as well: using the sigmoid threshold function and introducing low time lag resulted in a delay in the convergence by five iterations, while using the hyperbolic tangent function and introducing both the low time lag and a slightly higher one resulted in a delay in the convergence by four and six iterations respectively.

It should be noted, though, that iterations in a Fuzzy Cognitive Map simulation and real time units are not the same thing, and results should not be directly translated into time; this, however, can be partly overcome when causal relations are meaningful in the same time scale, as suggested by Kok [38]. In the same way that FCM results should not be used as real-value estimations [6, 52] but provide qualitative insight into the effectiveness of a set of policy strategies, the rate of causal propagation and eventually convergence too can provide qualitative insight into the rate of transition, without necessarily translating each simulation step into a time unit.

Given the need to limit global temperature increase within a strictly defined time frame in order to avoid the most significant and irreversible impacts, the introduction of the notion of time delays is of vital importance from a climate mitigation policy perspective.

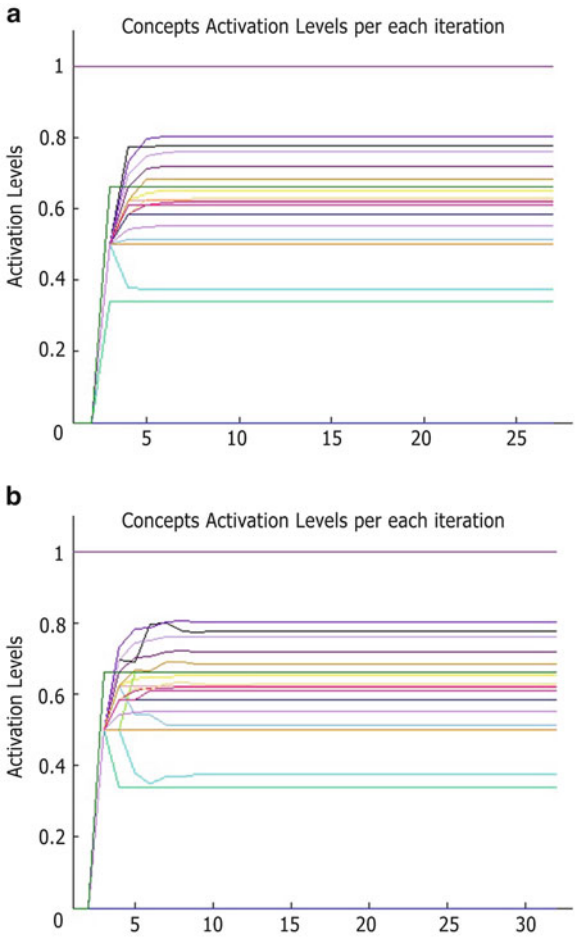


Fig. 11.5: Concept activation levels of the example FCM, using the sigmoid threshold function, (**a**) without time lag and (**b**) with slight time lag

11.5 Conclusions

The proposed methodological framework for selecting effective climate policies for low carbon transitions is based on the Fuzzy Cognitive Mapping method, and builds on the output of previously run quantitative models, an extensive literature review of potential risks and uncertainties, as well as country case studies of technological innovation systems, as part of a complete, analytical framework for developing robust transition pathways. This chapter, therefore, aims to present an innovative approach that takes advantage of the qualitative insight that Fuzzy Cognitive Maps can offer, suggesting however that the FCM technique cannot stand on its own when exploring alternative climate policy strategies but rather supplement existing quantitative methodologies, by linking them to experience-driven modeling.

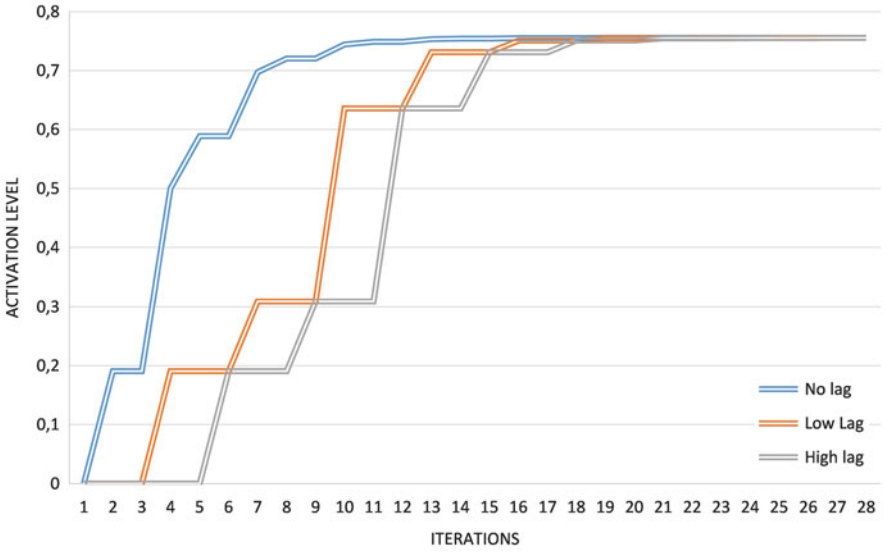


Fig. 11.6: Value increase in the goal concept, due to various time delays, using the hyperbolic tangent threshold function

Another key aspect of the presented approach lies in the simplification of the FCM model construction, from the experts' perspective. Although the required qualitative input is extended in order to include time-related issues, stakeholder engagement is limited to a clearly-defined and structured three-stage process. In this direction, extensive literature review on the various approaches of stakeholder engagement in fuzzy cognitive maps has been conducted, and specific research questions and steps have been determined, so as to reduce the method's complexity while at the same time not compromising the added value of the derived conclusions.

Last but not least, the proposed method introduces an ill-defined aspect of FCMs, which is the notion of time. By including causal relationship impacts that are defined in the same time unit, to the extent experts can offer such knowledge, and integrating potential time delays of these impacts into the model, the fuzzy cognitive maps of the examined systems become more dynamic, and new insight can be explored and gained with regard to each considered alternative, in an attempt to enhance the robustness of the resulting climate policy strategies.

Acknowledgements The most important part of this research is based on the H2020 European Commission Project "Transitions pathways and risk analysis for climate change mitigation and adaptation strategies—TRANSrisk" under grant agreement No. 642260. The sole responsibility for the content of this chapter lies with the authors. It does not necessarily reflect the opinion of the European Commission.

References

1. AdaptCost: Briefing paper 1: Integrated Assessment Models – Africa Results. Stockholm Environment Institute (2010)
2. Amer, M., Jetter, A., Daim, T.: Development of fuzzy cognitive map (FCM)-based scenarios for wind energy. Int. J. Energy Sect. Manage. **5**(4), 564–584 (2011)
3. Axelrod, R.: The analysis of cognitive maps. In: Axelrod, R. (ed.) Structure of Decision: The Cognitive Maps of Political Elites, pp. 55–73. Princeton University Press, Princeton (1976)
4. Barker, T.: Achieving a 10 % cut in Europe's carbon dioxide emissions using additional excise duties: coordinated, uncoordinated and unilateral action using the econometric model E3ME. Econ. Syst. Res. **11**(4), 401–422 (1999)
5. Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., Rickne, A.: Analyzing the functional dynamics of technological innovation systems: a scheme of analysis. Res. Policy **37**(3), 407–429 (2008)
6. Biloslavo, R., Dolinšek, S.: Scenario planning for climate strategies development by integrating group Delphi, AHP and dynamic fuzzy cognitive maps. Foresight **12**(2), 38–48 (2010)
7. Biloslavo, R., Grebenc, A.: Integrating group Delphi, analytic hierarchy process and dynamic fuzzy cognitive maps for a climate warning scenario. Kybernetes **41**(3/4), 414–428 (2012)
8. Bontis, N., Fitz-Enz, J.: Intellectual capital ROI: a causal map of human capital antecedents and consequents. J. Intellect. Cap. **3**(3), 223–247 (2002)
9. Bougon, M., Weick, K., Binkhorst, D.: Cognition in organizations: an analysis of the Utrecht jazz orchestra. Adm. Sci. Q. **22**(4), 606–639 (1977)
10. Bourgani, E., Stylios, C.D., Manis, G., Georgopoulos, V.C.: Integrated approach for developing timed fuzzy cognitive maps. In: Intelligent Systems' 2014, pp. 193–204. Springer, Berlin (2015)
11. Brown, S.R.: Political Subjectivity: Applications of Q Methodology in Political Science. Yale University Press, New Haven, Connecticut (1980)
12. Brown, S.M.: Cognitive mapping and repertory grids for qualitative survey research: some comparative observations. J. Manag. Stud. **29**(3), 287–307 (1992)
13. Bukowski, M., Kowal, P.: Large Scale, Multi-Sector DSGE Model as a Climate Policy Assessment Tool. Instytut Badań Strukturalnych, Warszawa (2010)
14. Bukowski, M., Kowal, P.: Modelling the transition to a low carbon economy with a large scale dsge model of Poland. Technical Report, IBS Technical Paper, Institute for Structural Research, Warsaw (2010)
15. Carvalho, J.P., Tomè, J.A.: Rule based fuzzy cognitive maps-expressing time in qualitative system dynamics. In: The 10th IEEE International Conference on Fuzzy Systems, vol. 1, pp. 280–283. IEEE (2001)
16. Commission, E.: A roadmap for moving to a competitive low carbon economy in 2050 (2011). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee, and the Committee of Regions, Document COM (2011): 112, http://eur-lex.europa.eu/legal--content/EN/TXT/PDF/?uri=CELEX:52011DC0112&from=en
17. Council, E.: European Council, 4 February 2011, Conclusions (2011). http://register.consilium.europa.eu/doc/srv?l=EN&f=ST%202%202011%20REV%201
18. Dalkey, N., Helmer, O.: An experimental application of the Delphi method to the use of experts. Manag. Sci. **9**(3), 458–467 (1963)
19. Dickerson, J., Kosko, B., et al.: Virtual worlds as fuzzy cognitive maps. In: IEEE Virtual Reality Annual International Symposium, pp. 471–477. IEEE (1993)
20. Eden, C., Ackermann, F.: Strategy Making: The Journey of Strategic Management. Sage, London (1998)
21. Eden, C., Ackermann, F., Cropper, S.: The analysis of cause maps. J. Manag. Stud. **29**(3), 309–324 (1992)

22. Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., et al.: Climate change 2014: mitigation of climate change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 511–597 (2014)
23. Edmonds, J., Wise, M., MacCracken, C.: Advanced energy technologies and climate change: an analysis using the global change assessment model (GCAM). Fondazione ENI Enrico Mattei (1994)
24. Fuss, S., Szolgayova, J., Obersteiner, M., Gusti, M.: Investment under market and climate policy uncertainty. Appl. Energy **85**(8), 708–721 (2008)
25. Groumpos, P.P.: Fuzzy cognitive maps: basic theories and their application to complex systems. In: Glykas, M. (ed.) Fuzzy Cognitive Maps, pp. 1–22. Springer, Berlin/ Heidelberg (2010)
26. Hagiwara, M.: Extended fuzzy cognitive maps. In: IEEE International Conference on Fuzzy Systems, pp. 795–801. IEEE (1992)
27. Harary, F., Norman, R., Cartwright, D.: Structural Models: An Introduction to the Theory of Directed Graphs. Wiley, New York (1965)
28. Hare, W.: A Safe Landing for the Climate. State of the World, pp. 13–29. Worldwatch Institute, New York (2009)
29. Heaps, C.: Long-Range Energy Alternatives Planning (leap) System [software version 2011.0043]. Stockholm Environment Institute, Somerville (2012)
30. Hobbs, B.F., Ludsin, S.A., Knight, R.L., Ryan, P.A., Biberhofer, J., Ciborowski, J.J.: Fuzzy cognitive mapping as a tool to define management objectives for complex ecosystems. Ecol. Appl. **12**(5), 1548–1565 (2002)
31. Huff, A.S.: Mapping Strategic Thought. Wiley, New York (1990)
32. Isak, K., Wildenberg, M., Adamescu, M., Skov, F., De Blust, G., Varjopuro, R.: Manual for applying fuzzy cognitive mapping–experiences from ALTER-Net. Alter-Net final report (2009)
33. Jetter, A.J.: Fuzzy cognitive maps for engineering and technology management: what works in practice? In: Technology Management for the Global Future. PICMET 2006, vol. 2, pp. 498–512. IEEE (2006)
34. Kabashi, S., Bekteshi, S., Ahmetaj, S., Kabashi, G., Najdovski, D., Zidanšek, A., Šlaus, I.: Effects of kosovo's energy use scenarios and associated gas emissions on its climate change and sustainable development. Appl. Energy **88**(2), 473–478 (2011)
35. Kandasamy, W., Smarandache, F.: Fuzzy Cognitive Maps and Neutrosophic Cognitive Map. Infinite Study, Phoenix (2004)
36. Kardaras, D., Karakostas, B.: The use of fuzzy cognitive maps to simulate the information systems strategic planning process. Inf. Softw. Technol. **41**(4), 197–210 (1999)
37. Kelly, D.L., Kolstad, C.D.: Integrated assessment models for climate change control. In: Folmer, H., Tietenberg, T. (eds.) International Yearbook of Environmental and Resource Economics 1999/2000, pp. 171–197. Edward Elgar, Cheltenham (1999)
38. Kok, K.: The potential of fuzzy cognitive maps for semi-quantitative scenario development, with an example from Brazil. Glob. Environ. Chang. **19**(1), 122–133 (2009)
39. Kosko, B.: Fuzzy cognitive maps. Int. J. Man Mach. Stud. **24**(1), 65–75 (1986)
40. Kulmer, V., Koland, O., Steininger, K.W., Fürst, B., Käfer, A.: The interaction of spatial planning and transport policy: a regional perspective on sprawl. J. Transport Land Use **7**(1), 57–77 (2014)
41. Laurenti, R., Lazarevic, D., Poulikidou, S., Montrucchio, V., Frostell, B., Bistagnino, L., Wennersten, R.: Using causal maps to identify potential sources of environmental impact outside the scope of LCA studies: preliminary findings from case studies on washing machines and road vehicles. In: 18th Annual International Sustainable Development Research Conference, 24–26 June 2012 (2012)
42. Lazarus, M., Heaps, C., Raskin, P.: Long range energy alternatives planning system (LEAP). Reference Manual. Boston Stockholm Environment Institute (SEI), Boston (1995)
43. Lee, K.C., Kim, J.S., Chung, N.H., Kwon, S.J.: Fuzzy cognitive map approach to web-mining inference amplification. Expert Syst. Appl. **22**(3), 197–211 (2002)

44. Lee, K.C., Lee, H., Lee, N., Lim, J.: An agent-based fuzzy cognitive map approach to the strategic marketing planning for industrial firms. Ind. Mark. Manag. **42**(4), 552–563 (2013)
45. Lopolito, A., Nardone, G., Prosperi, M., Sisto, R., Stasi, A.: Modeling the bio-refinery industry in rural areas: a participatory approach for policy options comparison. Ecol. Econ. **72**, 18–27 (2011)
46. Nordhaus, W.D.: The 'dice' model: background and structure of a dynamic integrated climate-economy model of the economics of global warming. Technical Report, Cowles Foundation for Research in Economics, Yale University (1992)
47. OECD: The Economic Consequences of Climate Change. OECD Publishing, Paris (2015)
48. Olazabal, M.: Sustainable urban development under climate change and resource scarcity. Ph.D. thesis, University of Cambridge (2015)
49. Ortiz, R.A., Golub, A., Lugovoy, O., Markandya, A., Wang, J.: The DICER model: methodological issues and initial results. Technical Report, Basque Centre for Climate Change/Klima Aldaketa Ikergai (2010)
50. Ozesmi, U.: Ecosystems in the mind: fuzzy cognitive maps of the kizilirmak delta wetlands in turkey. Ph.D. thesis, University of Minnesota (1999)
51. Özesmi, U., Özesmi, S.L.: Ecological models based on people's knowledge: a multi-step fuzzy cognitive mapping approach. Ecol. Model. **176**(1), 43–64 (2004)
52. Papageorgiou, E., Kontogianni, A.: Using fuzzy cognitive mapping in environmental decision making and management: a methodological primer and an application. In: Young, S.S., Silvern, S.E. (eds.) International Perspectives on Global Environmental Change, pp. 427–450. INTECH, Rijeka (2012)
53. Papageorgiou, E., Spyridonos, P., Glotsos, D.T., Stylios, C., Ravazoula, P., Nikiforidis, G., Groumpos, P.: Brain tumor characterization using the soft computing technique of fuzzy cognitive maps. Appl. Soft Comput. **8**(1), 820–828 (2008). doi:10.1016/j.asoc.2007.06.006. http://dx.doi.org/10.1016/j.asoc.2007.06.006
54. Papageorgiou, E., Markinos, A., Gemtos, T.: Soft computing technique of fuzzy cognitive maps to connect yield defining parameters with yield in cotton crop production in central Greece as a basis for a decision support system for precision agriculture application. In: Glykas, M. (ed.) Fuzzy Cognitive Maps, pp. 325–362. Springer, Berlin/Heidelberg (2010)
55. Papaioannou, M., Neocleous, C., Papageorgiou, C., Schizas, C.: A fuzzy cognitive map model of the repercussions of Greek PSI on Cypriot economy. In: 14th IEEE International Symposium on Computational Intelligence and Informatics (CINTI), pp. 255–260. IEEE (2013)
56. Park, K.S., Kim, S.H.: Fuzzy cognitive maps considering time relationships. Int. J. Hum. Comput. Stud. **42**(2), 157–168 (1995). doi:10.1006/ijhc.1995.1007. http://dx.doi.org/10.1006/ijhc.1995.1007
57. Roberts, F.: The questionnaire method. In: Axelrod, R. (ed.) Structure of Decision: The Cognitive Maps of Political Elites, pp. 333–342. Princeton University Press, Princeton (1976)
58. Rode, J., Crassard, F.: The Evolution of Building Integrated Photovoltaics (BIPV) in the German and French Technological Innovation Systems for Solar Cells. Chalmers University of Technology, Gothenburg (2007)
59. Scasny, M., Pisa, V., Pollit, H., Chewpreecha, U.: Analyzing macroeconomic effects of environmental taxation in the Czech Republic with the econometric E3ME model. Czech J. Econ. Finance **59**(5), 460–491 (2009)
60. Schneider, M., Shnaider, E., Kandel, A., Chew, G.: Automatic construction of FCMs. Fuzzy Sets Syst. **93**(2), 161–172 (1998). doi:10.1016/s0165-0114(96)00218-7. http://dx.doi.org/10.1016/S0165-0114(96)00218-7
61. Smith, J., Schellnhuber, H., Mirza, M.: Vulnerability to climate change and reasons for concern: a synthesis. In: McCarthy, J.J., White, K.S., Canziani, O., Leary, N., Dokken, D.J. (eds.) Climate Change 2001: Impacts: Adaptation and Vulnerability, pp. 913–967. Cambridge University Press, Cambridge (2001)
62. Stach, W., Kurgan, L., Pedrycz, W.: Expert-based and computational methods for developing fuzzy cognitive maps. In: Glykas, M. (ed.) Fuzzy Cognitive Maps, pp. 325–362. Springer, Berlin, Heidelberg (2010)

63. Stocker, T., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M.: Climate Change 2013: The Physical Science Basis. Cambridge University Press, Cambridge/New York (2014)
64. Styblinski, M., Meyer, B.: Signal flow graphs vs fuzzy cognitive maps in application to qualitative circuit analysis. Int. J. Man Mach. Stud. **35**(2), 175–186 (1991)
65. Stylios, C.D., Groumpos, P.P.: Modeling complex systems using fuzzy cognitive maps. IEEE Trans. Syst. Man Cybern. A Syst. Hum. **34**(1), 155–162 (2004)
66. Thomas, S.: The Analytic Hierarchy Process - Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York (1980)
67. UNFCCC: Report of the Conference of the Parties on Its Sixteenth Session, Held in Cancun from 29 November to 10 December 2010. FCCC/CP/2010/7/Add. 1, 15 March 2011, http://unfccc.int/files/na/application/pdf/07a01--1.pdf
68. UNFCCC: Adoption of the Paris Agreement, Proposal by the President, Draft Decision –/cp.21. Conference of the Parties, 21st Session (2015). http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf
69. van Vliet, M., Kok, K., Veldkamp, T.: Linking stakeholders and modellers in scenario studies: the use of fuzzy cognitive maps as a communication and learning tool. Futures **42**(1), 1–14 (2010)
70. Van Vuuren, D.P., Den Elzen, M.G., Lucas, P.L., Eickhout, B., Strengers, B.J., van Ruijven, B., Wonink, S., van Houdt, R.: Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs. Clim. Chang. **81**(2), 119–159 (2007)
71. Vidican, G., Samulewicz, D., McElvaney, L.: Challenges to developing technological innovation systems energy in Abu Dhabi. In: The 8th Triple Helix Conference, Madrid (2011)
72. Wang, K., Wang, C., Chen, J.: Analysis of the economic impact of different Chinese climate policy options based on a CGE model incorporating endogenous technological change. Energy Policy **37**(8), 2930–2940 (2009)
73. Watkiss, P., Downing, T., Dyszynski, J.: AdaptCost Project: Analysis of the Economic Costs of Climate Change Adaptation in Africa. UNEP, Nairobi (2010)
74. Wei, Z., Lu, L., Yanchun, Z.: Using fuzzy cognitive time maps for modeling and evaluating trust dynamics in the virtual enterprises. Expert Syst. Appl. **35**(4), 1583–1592 (2008)
75. Weyant, J., Davidson, O., Dowlabathi, H., Edmonds, J., Grubb, M., Parson, E., Cline, W.: Integrated assessment of climate change: an overview and comparison of approaches and results. In: Bruce, J., Lee, H., Haites, E. (eds.) Climate Change 1995: Economic and Social Dimensions of Climate, pp. 367–396. Cambridge University Press, Cambridge (1996)
76. Wiseman, J., Edwards, T., Luckins, K.: Post carbon pathways: A meta-analysis of 18 large-scale post carbon economy transition strategies. Environ. Innov. Soc. Transit. **8**, 76–93 (2013)
77. Wrightson, M.: The documentary coding method. In: Axelrod, R. (ed.) Structure of Decision: The Cognitive Maps of Political Elites, pp. 291–332. Princeton University Press, Princeton (1976)
78. Xirogiannis, G., Glykas, M.: Fuzzy cognitive maps in business analysis and performance-driven change. IEEE Trans. Eng. Manag. **51**(3), 334–351 (2004)

Chapter 12
Robust Optimization Approaches to Single Period Portfolio Allocation Problem

Nalân Gülpınar and Zhezhi Hu

Abstract Portfolio management is one of the fundamental problems in financial decision making. In a typical portfolio management problem, an investor is concerned with an optimal allocation of the capital among a number of available financial assets to maximize the return on the investment while minimizing the risk. This problem was formulated in the mean-variance portfolio management framework proposed by Markowitz in 1952. Since then, it has been widely studied by researchers and the practitioners. However, the solution is sensitive to model parameters due to data uncertainty. In this chapter, we review robust approaches to deal with data uncertainty for a single-period portfolio allocation problem. We first introduce the main ideas of robust optimization using symmetric and asymmetric uncertainty sets where the uncertain asset returns belong to. We then focus on data driven and distributionally robust optimization approaches.

12.1 Introduction

In financial portfolio management, a rational framework for investment decisions is provided by the maximization of expected portfolio return for an acceptable level of risk. The single-period mean-variance portfolio management framework is a fundamental example introduced by Markowitz [51]. Since then, various stochastic programming approaches have been developed to solve the underlying optimization problem.

The single-stage model of Markowitz considers a portfolio to be constructed by M risky assets with random rates of return $\tilde{r}_1, \cdots, \tilde{r}_M$. Note that vectors and matrices

N. Gülpınar (✉) • Z. Hu
Warwick Business School, The University of Warwick, Coventry, UK
e-mail: Nalan.Gulpinar@wbs.ac.uk; zhezhi.hu.14@mail.wbs.ac.uk

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_12

are represented in bold and prime $(*)'$ refers to transpose operation. For instance, $\tilde{\mathbf{r}} = (\tilde{r}_1, \cdots, \tilde{r}_M)'$ is a vector of rates of returns. Let $\mathbf{x} = (x_1, \cdots, x_M)'$ represent the vector of asset weights. We assume that short-selling is not allowed; thus, $\mathbf{x} \geq 0$.

The initial budget is normalized to 1; that is $\mathbf{e}'\mathbf{x} = 1$ where $\mathbf{e} = (1, \cdots, 1)'$. The portfolio return is computed as $\tilde{r}_p = \tilde{\mathbf{r}}'\mathbf{x}$.

Given a probability distribution $\mathbb{P}$ of random rates of the asset returns $\tilde{\mathbf{r}}$, one can compute the expected portfolio return $\boldsymbol{\mu} = \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{r}}]$ and expected portfolio risk that is defined as the variance of portfolio return, $\boldsymbol{\Sigma} = \mathbb{E}_{\mathbb{P}}[(\tilde{\mathbf{r}} - \boldsymbol{\mu})'(\tilde{\mathbf{r}} - \boldsymbol{\mu})]$. Then the single stage mean-variance portfolio allocation problem is equivalent to the following deterministic convex quadratic programming problem:

$$\max_{\mathbf{x} \in X} \quad \lambda \boldsymbol{\mu}'\mathbf{x} - (1 - \lambda)\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x}, \tag{12.1}$$

where $\mathbf{X} = \{\mathbf{x} \mid \mathbf{e}'\mathbf{x} = 1,\ \mathbf{x} \geq 0\}$ is the feasible region.

The scaling constant $\lambda \in [0, 1]$ determines the level of risk aversion. By varying λ from 0 (total risk aversion) to 1 (total risk-seeking), the investor can adjust the trade-off between the portfolio return and risk preferred. Thus, the entire range of efficient investment strategies is determined. The efficient frontier displays all possible portfolios in risk-return space.

Although the mean-variance portfolio allocation model has been widely used by both academics and practitioners, it has been criticized for several reasons. One of the main drawbacks is that the optimal asset allocation is very sensitive to model parameters. The true distribution of the asset returns is not known precisely, but has to be estimated from the available data, which might be insufficient to identify the unique distribution. Imprecise forecasts and inaccurate estimated model parameters involve some errors. Such estimation errors have serious consequences in the optimal investment portfolio obtained from the model; for instance see [48, 52]. The empirical studies showed that the portfolio weights were extremely sensitive to the estimation errors in mean asset returns [15, 22]. Moreover, Ceria and Stubbs [19] argued that the mean-variance model has the error-maximization property. Bertsimas et al. [13] also debated that such parameter uncertainty may result in suboptimal and infeasible solution since the solution quality highly relies on the estimation accuracy of these parameters.

Stochastic programming models describe underlying uncertainties in financial optimization problems in view of expected value decision criteria. They assume that the uncertain parameters in the optimization problems follow a known distribution. There are different methods to deal with uncertain data such as scenario-based stochastic programming and chance-constrained optimization. A scenario-based stochastic programming approach takes into account a finite number of realizations of the random variables and specifies the optimal decisions in view of these scenarios [23]. Chance-constrained stochastic programming involves probabilistic constraints to control risk in decision making under uncertainty. The reader is referred to [16, 46, 58] for more information on other applications of stochastic programming in different sectors.

It is worthwhile to mention that there are alternative stochastic optimization techniques based on dynamic programming algorithms that require specific modeling skills using states and actions that correspond to random paths and decisions in the multistage stochastic programming setting. However, these models also suffer from the curse of dimensionality in the state and action spaces. To deal with this, simulation-based dynamic programming approaches have been developed to solve the underlying problem approximately using forward dynamic programming algorithms. They have also been successfully applied to real life applications. The reader is referred to [57] for an overview and various applications of approximate dynamic programming.

Worst-case analysis is used as an alternative approach to protect against the risk of inaccurate estimations. Rustem et al. [61] and Gulpinar and Rustem [38] considered many possible rival return scenarios and covariance matrices within mean-variance portfolio management framework. They applied a min-max framework to find the best action (i.e. investment portfolio) under the worst-case rival scenarios. It is shown that the optimal investment decision in view of different rival scenarios is robust to the parameter uncertainty. While the multiple rival estimates can be regarded as a discretized description of the parameter uncertainty, a more general approach in the continuous case is to assume the random variable belongs to an uncertainty set. Such uncertainty set usually contains an infinite number of possible realizations around the estimated value of the random variable and can be regarded as the decision maker's confidence about the estimated value. This approach is referred to as *robust optimization* that is the main subject of this chapter. Moreover, since the true distribution of the random variable is not known exactly, one can also consider a distribution set to describe the uncertainty of the underlying probability distribution. This approach is referred to as *distributionally robust optimization*, that will be also covered in this chapter. We aim to review different robust optimization approaches in the context of portfolio management problems.

The remaining of this chapter is organized as follows. We describe the robust portfolio management model in Sect. 12.2. Section 12.3 focuses on how to define an uncertainty set. The robust counterparts of the portfolio allocation problem are then derived in view of symmetric and asymmetric uncertainty sets in Sect. 12.4. Based on recent developments in data-driven robust optimization, we show how to construct data-driven uncertainty sets in Sect. 12.5. Section 12.6 introduces the application of distributionally robust optimization for robust portfolio management problems. We also present the connections between robust optimization and risk measures in Sect. 12.7. A brief summary of the chapter and future research directions are presented in Sect. 12.8.

12.2 Robust Portfolio Management Model

Robust optimization deals with data uncertainty using the worst-case analysis for a pre-specified uncertainty set. The use of uncertainty set in describing parameter uncertainty was first suggested by Soyster [22]. After being neglected for a

while, robust optimization was extended by Ghaoui and Lebret [16] and Ben-Tal and Nemirovski [4]. Since then, it has been applied for solving various practical problems in different areas. The robust optimization approach assumes that the uncertain parameter arising in the optimization problem belongs to an uncertainty set that can be constructed from the probability distributions of uncertain factors. A robust counterpart of the original problem is then derived in such a way that the optimal solution to the optimization problem remains feasible for all realizations of the stochastic data within the pre-specified uncertainty sets, including the worst-case values. Depending on the specification of the uncertainty sets, the robust counterpart of the original optimization problem can be formulated as a tractable optimization problem with no random parameter. For further information on robust optimization and recent developments, the reader is referred to [8].

In terms of financial applications, robust optimization has been mainly used for the portfolio management problems. Ben-Tal and Nemirovski [5] provided tractable robust formulations for a single-period portfolio management problem using an ellipsoidal uncertainty set. Goldfarb and Iyengar [36] developed a robust factor model and showed how to derive the uncertainty set by statistical procedures. Ghaoui et al. [33] introduced tractable formulation for worst-case value-at-risk in robust portfolio optimization. Tutuncu and Koenig [64] used interval uncertainty sets for robust mean-variance model. Natarajan et al. [53] studied a robust portfolio optimization problem using asymmetric distributional information in random variables. Zhu and Fukushima [66] used worst-case conditional value-at-risk in robust portfolio optimization under box and ellipsoidal uncertainty sets. Delage and Ye [24] developed a distributional robust approach for single-period portfolio management problem. Gulpinar et al. [39] considered robust portfolio allocation problem in view of discrete asset choice constraints.

Mathematically, the robust optimization provides a different description of the portfolio management problem compared to the classical stochastic mean-variance model. It assumes that the uncertain asset returns $\tilde{\mathbf{r}}$ belong to an uncertainty set $\mathscr{U}_{\tilde{\mathbf{r}}}$ Instead of maximizing the expected portfolio return $\mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{r}}]$, we optimize the worst-case portfolio return in view of the uncertainty set $\mathscr{U}_{\tilde{\mathbf{r}}}$. In other words, an investor is concerned with a robust investment decision $\mathbf{x}$ such that the worst-case portfolio return is maximized. This can be formulated as the following max-min problem:

$$\max_{\mathbf{x}\in\mathbf{X}} \min_{\tilde{\mathbf{r}}\in\mathscr{U}_{\tilde{\mathbf{r}}}} \tilde{\mathbf{r}}'\mathbf{x}. \tag{12.2}$$

We now introduce an auxiliary variable γ, and rewrite (12.2) as:

$$\begin{aligned} &\max_{\mathbf{x}\in\mathbf{X},\gamma} \quad \gamma \\ &s.t. \quad \min_{\tilde{\mathbf{r}}\in\mathscr{U}_{\tilde{\mathbf{r}}}} \tilde{\mathbf{r}}'\mathbf{x} \geq \gamma. \end{aligned} \tag{12.3}$$

The choice of uncertainty sets plays an important role on the investment decisions. Once the uncertainty sets are specified, the inner minimization problem is solved to derive the corresponding robust counterpart. The robust counterpart of the linear constraint in view of the chosen uncertainty set leads a tractable deterministic problem.

12.3 Defining Uncertainty Sets

Different uncertainty sets lead to different robust counterparts of the original constraint with different computational burdens. Bertsimas et al. [14] described two main criteria to be satisfied when we define an uncertainty set $\mathscr{U}_{\tilde{\mathbf{r}}}$:

- $\min_{\tilde{\mathbf{r}} \in \mathscr{U}_{\tilde{\mathbf{r}}}} \tilde{\mathbf{r}}'\mathbf{x} \geq \gamma$ is computationally tractable.
- If $\min_{\tilde{\mathbf{r}} \in \mathscr{U}_{\tilde{\mathbf{r}}}} \tilde{\mathbf{r}}'\mathbf{x} \geq \gamma$, then $Prob(\tilde{\mathbf{r}}'\mathbf{x} \geq \gamma) \geq 1 - \varepsilon$.

The first criterion ensures the computational tractability of the original optimization problem that depends on the type of the uncertainty sets. There are various symmetric and asymmetric uncertainty sets that have been used for various applications of robust optimization. The shape of the uncertainty set defines a risk measure on the constraints with uncertain coefficients [54]. In practice, the shape is selected to reflect the modeler's knowledge of the probability distributions of the uncertain parameters, keeping in mind that, ideally, the robust counterpart problem should be efficiently solvable if the uncertainties are assumed to belong to that uncertainty set. The ellipsoidal uncertainty set, for example, defines a standard-deviation-like risk measure on the constraint with uncertain parameters, and in the case of linear optimization, results in a robust counterpart to the original problem that is a second order cone problem (that is a tractable optimization problem).

When solving optimization problems with uncertain parameters using the robust optimization approach, the size of the specified uncertainty set is often related to guarantees on the probability that the constraint with uncertain coefficients will not be violated (see, for example, [12]). The second criterion shows that for given probability $\varepsilon \geq 0$, the uncertainty set $\mathscr{U}_{\tilde{\mathbf{r}}}$ should imply a probability guarantee for the true distribution $\mathbb{P}^*$ of random variable $\tilde{\mathbf{r}}$. If we seek for a higher probability guarantee for $\mathbb{P}^*$ (i.e. ε is closer to 0), then $\mathscr{U}_{\tilde{\mathbf{r}}}$ should be larger to include more possible realizations of $\tilde{\mathbf{r}}$. Since we then optimize against the worst-case value within $\mathscr{U}_{\tilde{\mathbf{r}}}$, the robust solution is more conservative. Bertsimas and Sim [12] showed that there is always a trade-off between solution optimality and conservativeness when choosing uncertainty sets. By changing the size of the uncertainty set, one can flexibly adjust the level of conservativeness of robust solutions according to the investor's preference. Furthermore, recent development in data-driven optimization has shown that one can construct a less conservative uncertainty set by incorporating additional information from historical data, where such uncertainty set can still imply the same probability guarantee as before. These issues will be explained further in details in Sect. 12.5.

12.4 Derivation of Robust Counterpart

Next, we derive the robust counterpart of the portfolio allocation model using different discrete and continuous uncertainty sets and discuss their computational burdens. The discrete uncertainty set involves a finite number of risk and return

realizations while the continuous sets consist of an infinite number of future observations with symmetric and asymmetric structures. The symmetric and asymmetric shapes of the uncertainty sets should allow us to map the uncertainty sets better to uncertain parameters with symmetric and skewed distributions, respectively.

Scenario-Based Uncertainty Sets Suppose we have generated different return scenarios and estimated different covariance matrices. Let I and K be a number of rival return and risk scenarios, respectively. A discrete uncertainty set is

$$\mathscr{U}_d = \{\tilde{\mathbf{r}} \in \{\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_I}\},\ \tilde{\Sigma} \in \{\Sigma_1, \cdots, \Sigma_K\}\}.$$

A robust mean-variance portfolio allocation problem can be formulated in a min-max compact form as follows:

$$\min_{\mathbf{x}\in\mathbf{X}} \left\{ (1-\lambda) \max_{k=1,\cdots,K} \mathbf{x}'\Sigma_k\mathbf{x} - \lambda \min_{i=1,\cdots,I} \mathbf{x}'\mathbf{r_i} \right\}.$$

In order to find the worst-case investment decision, we introduce free decision variables ν_1 and ν_2 to represent the inner optimization problems. The corresponding robust formulation of the problem can be reformulated as a quadratically constrained mathematical problem:

$$\begin{array}{lll} \min\limits_{\mathbf{x}\in\mathbf{X},\nu_1,\nu_2} & (1-\lambda)\nu_1 - \lambda\nu_2 & \\ s.t. & \mathbf{x}'\Sigma_k\mathbf{x} \le \nu_1, & k = 1,\cdots,K \\ & \mathbf{x}'\mathbf{r_i} \ge \nu_2, & i = 1,\cdots,I. \end{array}$$

Symmetric Uncertainty Sets Consider symmetric (ellipsoidal) uncertainty sets involving the uncertain future asset returns $\tilde{r}_i$ for $i = 1,\cdots,M$. The uncertainty set is specified in terms of estimated mean vectors $\hat{\mu}$ and the covariance matrices $\hat{\Sigma}$ of the vectors of random variables $\tilde{\mathbf{r}}$ in the set of constraints with uncertain coefficients as follows:

$$\mathscr{U}_{\tilde{\mathbf{r}}} = \{\tilde{\mathbf{r}} : \|\hat{\Sigma}^{-\frac{1}{2}}(\tilde{\mathbf{r}} - \hat{\mu})\|_2 \le \Omega\},$$

where $\|\cdot\|_2$ represents the Euclidean norm and Ω is the budget of robustness predefined to adjust the size of ellipsoidal uncertainty set.

Ben-Tal and Nemirovski [6] and Bertsimas and Sim [12] showed that $\Omega = \sqrt{-2\ln\varepsilon}$ implies a probability guarantee $Prob(\tilde{\mathbf{r}}'\mathbf{x} \ge \gamma) \ge 1-\varepsilon$. If ε is close to 0, which means $\tilde{\mathbf{r}}'\mathbf{x} \ge \gamma$ should be guaranteed with a very high probability (hard constraint), then $\mathscr{U}_{\tilde{\mathbf{r}}}$ is chosen as extremely large set since Ω goes to infinity. In contrast, if $\varepsilon = 1$, then $\mathscr{U}_{\tilde{\mathbf{r}}}$ becomes a singleton $\{\hat{\mu}\}$ and the estimation error of $\hat{\mu}$ is ignored. In order to derive the robust counterpart of (12.2) in view of ellipsoidal uncertainty set, we need to solve the inner minimization problem $\min\limits_{\tilde{\mathbf{r}}\in\mathscr{U}_{\tilde{\mathbf{r}}}} \tilde{\mathbf{r}}'\mathbf{x}$ that can be rewritten as

$$\begin{array}{ll} \min\limits_{\tilde{\mathbf{r}}} & \mathbf{x}'\tilde{\mathbf{r}} \\ s.t. & \|\hat{\Sigma}^{-\frac{1}{2}}(\tilde{\mathbf{r}} - \hat{\mu})\|_2 \le \Omega. \end{array}$$

Let's introduce Lagrangian multiplier $\eta \geq 0$ for the constraint of the optimization problem. The Lagrangian function is:

$$\mathscr{L}(\tilde{\mathbf{r}}, \eta) = \mathbf{x}'\tilde{\mathbf{r}} + \eta \left(\|\hat{\Sigma}^{-\frac{1}{2}}(\tilde{\mathbf{r}} - \hat{\mu})\|_2 - \Omega \right).$$

Applying the first order optimality and complementarity conditions, we can find the optimal value of $\tilde{\mathbf{r}}$ within the uncertainty set $\mathscr{U}_{\tilde{\mathbf{r}}}$ as $\tilde{\mathbf{r}} = \hat{\mu} + \left(-\frac{\Omega}{\eta}\right)\hat{\Sigma}\mathbf{x}$. In addition, the optimal value of the Lagrangian multiplier is obtained as $\eta = \sqrt{\mathbf{x}'\hat{\Sigma}\mathbf{x}}$. The objective function value of the inner minimization problem becomes:

$$\mathbf{x}'\tilde{\mathbf{r}} = \hat{\mu}'\mathbf{x} - \Omega\sqrt{\mathbf{x}'\hat{\Sigma}\mathbf{x}}.$$

The robust counterpart of (12.2) can be derived as the following second order conic programming model:

$$\max_{\mathbf{x}\in\mathbf{X}} \quad \hat{\mu}'\mathbf{x} - \Omega\sqrt{\mathbf{x}'\hat{\Sigma}\mathbf{x}}. \tag{12.4}$$

Notice that (12.4) is quite similar to the mean-variance portfolio allocation problem in (12.1). Both models take into account the trade-off between portfolio return and portfolio risk. The main difference between these models in (12.1) and (12.4) is that portfolio covariance ($\mathbf{x}'\Sigma\mathbf{x}$) is replaced by its standard deviation ($\sqrt{\mathbf{x}'\Sigma\mathbf{x}}$) and the risk averse parameter λ is replaced by Ω. Zymler et al. [67] stated that for each λ in (12.1) there is one Ω such that (12.1) and (12.4) have the same optimal solution.

Asymmetric Uncertainty Sets A symmetric uncertainty set (such as ellipsoidal) can represent the underlying uncertainty well when it follows a symmetric probability distribution such as the normal distribution. Theoretically, the assumption that asset returns follow normal distributions is not unreasonable because the Central Limit Theorem implies that over the long horizon, returns should be approximately Gaussian as long as short-horizon returns are sufficiently independent (see, for example, [18]). Empirically, however, there is evidence that both short- and long-horizon stock returns can be skewed and highly leptokurtic (see, for example, [25, 30]). In fact, asset returns are asymmetrically distributed due to time dependence [27], market driven [29] and different structural features such as skewness and kurtosis [28, 50]. In these cases, using a symmetric uncertainty set may result in too conservative results.

In order to overcome the drawback of symmetric uncertainty set, Chen et al. [21] introduced forward and backward deviations to incorporate asymmetric distribution information of random variables when defining uncertainty sets. This idea was then applied in a single-period robust portfolio management problem by Natarajan et al. [53]. They assumed that random asset return is generated by the following factor model:

$$\tilde{\mathbf{r}} = \hat{\mu} + \hat{\Sigma}^{1/2}\tilde{\mathbf{z}},$$

where the vector $\tilde{\mathbf{z}}$ can be regarded as the *normalized* asset return which has zero mean and consists of independent factors. Moreover, the vector of random variables $\tilde{\mathbf{z}}$ is decomposed into two vectors of random variables $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{h}}$ such that $\tilde{\mathbf{z}} = \tilde{\mathbf{g}} - \tilde{\mathbf{h}}$ where $\tilde{\mathbf{g}} = \max\{\tilde{\mathbf{z}}, 0\}$ and $\tilde{\mathbf{h}} = \max\{-\tilde{\mathbf{z}}, 0\}$. Let $p_j > 0$ and $q_j > 0$ represent the forward and backward deviations of the random variable $\tilde{z}_j$ for $j = 1, \cdots, M$ respectively. Let's define diagonal matrices $\mathbf{P} = diag(p_1, \cdots, p_M)$ and $\mathbf{Q} = diag(q_1, \cdots, q_M)$. An asymmetric uncertainty set is described as:

$$\mathscr{U}_{\tilde{\mathbf{z}}} = \{\mathbf{z} : \exists \mathbf{g}, \mathbf{h} \in \mathbb{R}^M,\ \mathbf{z} = \mathbf{g} - \mathbf{h},\ \|\mathbf{P}^{-1}\mathbf{g} + \mathbf{Q}^{-1}\mathbf{h}\| \leq \Gamma,\ \underline{\mathbf{z}} \leq \mathbf{z} \leq \overline{\mathbf{z}}\},$$

where $\tilde{\mathbf{z}}$ is assumed to have a support as $[\underline{\mathbf{z}}, \overline{\mathbf{z}}]$ to ensure $p(\tilde{\mathbf{z}})$ and $q(\tilde{\mathbf{z}})$ to be finite. Note that Γ represents the desired degree of robustness. In the case of a symmetric distribution, the uncertainty set above is ellipsoidal and we have $p(\tilde{\mathbf{z}}) = q(\tilde{\mathbf{z}})$. In other words, this asymmetric uncertainty set includes the ellipsoidal uncertainty set as a special case. If the factors $\tilde{\mathbf{z}}$ are independent and Γ is selected so that $\Gamma \geq \sqrt{-2ln\varepsilon}$, then (as in the case of the ellipsoidal uncertainty set) the constraint will be satisfied with probability of at least $1 - \varepsilon$ [53].

For a general case, we have $p(\tilde{\mathbf{z}}) \neq q(\tilde{\mathbf{z}})$. Moreover, $\mathbf{P}^{-1}\mathbf{g}$ and $\mathbf{Q}^{-1}\mathbf{h}$ normalize the positive and negative deviations of $\mathbf{z}$ respectively, which are used to form an uncertainty set with budget Γ. Thus, $\mathscr{U}_{\tilde{\mathbf{z}}}$ actually consists of a box uncertainty set $\{\mathbf{z} : \underline{\mathbf{z}} \leq \mathbf{g} - \mathbf{h} \leq \overline{\mathbf{z}}\}$ and an ellipsoidal uncertainty set $\{\mathbf{z} : \|\mathbf{P}^{-1}\mathbf{g} + \mathbf{Q}^{-1}\mathbf{h}\| \leq \Gamma\}$. The robust counterpart of the portfolio management problem (12.2) using the asymmetric uncertainty set $\mathscr{U}_{\tilde{\mathbf{z}}}$ can be formulated as the following optimization problem:

$$\max_{\mathbf{x} \in \mathbf{X}} \ \tilde{\mu}'\mathbf{x} + \min_{\tilde{\mathbf{z}} \in \mathscr{U}_{\tilde{\mathbf{z}}}} (\hat{\Sigma}^{1/2}\tilde{\mathbf{z}})'\mathbf{x}.$$

The inner minimization problem can be rewritten as:

$$\begin{aligned} \min_{\mathbf{g},\mathbf{h}} \ & (\hat{\Sigma}^{1/2}\mathbf{x})'(\mathbf{g} - \mathbf{h}) \\ s.t. \ & \|\mathbf{P}^{-1}\mathbf{g} + \mathbf{Q}^{-1}\mathbf{h}\| \leq \Gamma \\ & \mathbf{g} - \mathbf{h} \leq \overline{\mathbf{z}} \\ & \mathbf{g} - \mathbf{h} \geq \underline{\mathbf{z}} \\ & \mathbf{g}, \mathbf{h} \geq 0. \end{aligned}$$

Using duality theory, the robust counterpart of (12.2) is derived as a second order conic programming model:

$$\begin{aligned} \max_{\mathbf{x} \in \mathbf{X}, \theta, \alpha, \beta, \delta} \ & \theta \\ s.t. \ & \hat{\mu}'\mathbf{x} + \theta \geq \Gamma\|\delta\| + \alpha'\overline{\mathbf{z}} + \beta'\underline{\mathbf{z}} \\ & \delta_i \geq p_i \left(\mathbf{e}_i'(\hat{\Sigma}^{1/2})'\mathbf{x} + \alpha_j - \beta_j\right), \quad i = 1, \ldots, M \\ & \delta_i \geq -q_i \left(\mathbf{e}_i'(\hat{\Sigma}^{1/2})'\mathbf{x} + \alpha_j - \beta_j\right), \ i = 1, \ldots, M \\ & \alpha, \beta \geq 0. \end{aligned}$$

So far, we have introduced symmetric and asymmetric uncertainty sets. All symmetric uncertainty sets are chosen based on some priori assumptions about certain structural features of $\mathbb{P}^*$. For example, one firstly assumes that the uncertainty set $\mathscr{U}$ is ellipsoidal. Then by estimating the first two moments of the random variable, the robust counterpart of the original problem is derived. Also, the budget of robustness, which is an important factor to adjust the size of the uncertainty set, is predefined.

Similarly, for asymmetric uncertainty sets, the historical data can be used to incorporate asymmetric distribution information (i.e. to estimate forward and backward deviation) about the random variable. But the structure of the uncertainty set is still based on priori assumption (i.e. an ellipsoid). Both Chen et al. [21] and Natarajan et al. [53] found that asymmetric sets can provide less conservative solutions than symmetric ones. These studies also showed the potential benefit of using data when building uncertainty sets, which leads to a unified approach called data-driven robust optimization.

12.5 Data-Driven Robust Optimization

In data-driven robust optimization methods, the uncertainty sets are constructed with certain structures and sizes that accord with the data. These sets still satisfy the two criteria introduced in Sect. 12.3. In order to make sure that the data-driven sets can imply a similar probability guarantee as non data-driven sets (i.e. those introduced in Sect. 12.4), it is natural to link them with confidence regions. For a detailed discussion on the use of data-driven robust optimization approaches to construct uncertainty sets based on confidence regions, the reader is referred to [14].

From a portfolio management perspective, investors may consider the historical data (asset returns) when making investment decisions. It is reasonable that such "big data" can help investors construct uncertainty sets and yield less conservative robust investment decisions. The use of data in constructing uncertainty sets has been proposed in several studies. Goldfarb and Iyengar [36] considered market data to estimate parameters of a return factor model through linear multivariate regression. They built confidence regions around these estimated parameters and used them to construct uncertainty sets for any desired confidence level. Ben-Tal et al. [9] discussed how to construct uncertainty sets using statistical confidence region based on ϕ-divergences. They showed the tractability of the robust counterpart using these uncertainty sets and applied them to an asset pricing problem. Bertsimas et al. [14] introduced several kinds of uncertainty sets constructed by different hypothesis tests and applied them for a single-stage robust portfolio management problem. The empirical results confirmed that the data-driven uncertainty sets perform better than that constructed by traditional techniques (no data driven).

Next we will briefly summarize how to construct a data-driven uncertainty set by a confidence region based on the work of Goldfarb and Iyengar [36]. We assume that the vector of asset returns $\tilde{\mathbf{r}}$ is generated by a vector autoregression VAR(1) process:

$$\tilde{\mathbf{r}}_t = \mathbf{c} + \mathbf{A}\tilde{\mathbf{r}}_{t-1} + \tilde{\varepsilon}_t,$$

where $\tilde{\varepsilon}_t$ represents the random residuals which is assumed to be independent among different time periods t and follows a multivariate normal distribution: $\tilde{\varepsilon}_t \sim \mathcal{N}(0, \sigma^2)$. More precisely, we can write VAR(1) process for each asset i at time t as:

$$\tilde{r}_{i,t} = c_i + \mathbf{a}_i \tilde{\mathbf{r}}_{t-1} + \tilde{\varepsilon}_{i,t}, \ for\ i = 1, \ldots, M,$$

where c_i is the i-th element of intercept vector $\mathbf{c}$ and $\mathbf{a}_i$ is the i-th row of matrix $\mathbf{A}$ that consists of the slope coefficients and $\tilde{\varepsilon}_{i,t} \sim \mathcal{N}(0, \sigma_i^2)$.

Given the historical data of all M asset returns for past N months, our goal is to construct an uncertainty set for the random asset return in the next months. Let $\mathbf{B} = [\mathbf{1}, \mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{N-1}] \in \mathbb{R}^{(N-1)\times(M+1)}$ and $\mathbf{y}_i = [r_{i,2}, r_{i,3}, \ldots, r_{i,N}] \in \mathbb{R}^{N-1}$ represent the known data parameters. We denote a vector of all coefficients in VAR(1) process as $\mathbf{w}_i = [c_i, \mathbf{a}_i] \in \mathbb{R}^{M+1}$.

For the vector of error terms $\tilde{\mathbf{u}}_i = [\tilde{\varepsilon}_{i,2}, \tilde{\varepsilon}_{i,3}, \cdots, \tilde{\varepsilon}_{i,N}] \in \mathbb{R}^{N-1}$, the VAR(1) process for each asset i becomes

$$\mathbf{y}_i = \mathbf{B}\mathbf{w}_i + \tilde{\mathbf{u}}_i, \ for\ i = 1, \ldots, M.$$

Goldfarb and Iyengar [36] showed that the model parameters can be estimated using the least squares method as follows:

$$\hat{\mathbf{w}}_i = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{y}_i, \ for\ i = 1, \cdots, M.$$

Moreover, the estimation error $\mathbf{w}_i$ follows a multivariate normal distribution:

$$\hat{\mathbf{w}}_i - \mathbf{w}_i = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\tilde{\mathbf{u}}_i \sim \mathcal{N}(0, \sigma_i^2(\mathbf{B}'\mathbf{B})^{-1})$$

that implies

$$\frac{1}{\sigma_i^2}(\hat{\mathbf{w}}_i - \mathbf{w}_i)'(\mathbf{B}'\mathbf{B})(\hat{\mathbf{w}}_i - \mathbf{w}_i) \sim \chi^2_{j+1},$$

where $rank(\mathbf{B}) = j+1$. Although the true residual variance σ_i^2 is not known, it can be estimated from historical data as:

$$\hat{\sigma}_i^2 = \frac{\|\mathbf{y}_i - \mathbf{B}\hat{\mathbf{w}}_i\|}{N-j-1}.$$

Using the relationship between χ^2 and F distributions (see for instance, [1, 37]), we can observe that the resulting random variable:

$$\frac{1}{(j+1)\hat{\sigma}_i^2}(\hat{\mathbf{w}}_i - \mathbf{w}_i)'(\mathbf{B}'\mathbf{B})(\hat{\mathbf{w}}_i - \mathbf{w}_i)$$

also follows a F distribution. Thus, for a given confidence level κ and critical value $c_{j+1}(\kappa)$, we can construct a confidence region of $\mathbf{w}_i$ as

$$\mathcal{U}_{\tilde{\mathbf{w}}_i} = \{\mathbf{w}_i : (\hat{\mathbf{w}}_i - \mathbf{w}_i)'(\mathbf{B}'\mathbf{B})(\hat{\mathbf{w}}_i - \mathbf{w}_i) \leq (j+1)c_{j+1}(\kappa)\hat{\sigma}_i^2\}.$$

Note that this can be regarded as an uncertainty set for the coefficients of VAR(1) process. For a non data-driven uncertainty sets (such as symmetric and asymmetric uncertainty sets introduced in previous section), the price of robustness Ω is set as a priori assumption. However, in the data-driven uncertainty set, the size is not only influenced by priori assumption (confidence level κ), but also the estimated residual variance $\hat{\sigma}_i^2$ from the data.

12.6 Distributionally Robust Optimization

In robust optimization, the uncertainty of random variables is described by deterministic sets rather than probability distributions as in stochastic programming. Worst-case values are chosen within the sets to imply a high probability guarantee that the solution is robust to randomness. From the previous section, we notice that by exploring historical data to get additional information about the probability distribution of random variables, we can construct the uncertainty sets that fit better with the data. So a reasonable extension to the robust optimization is to combine the ideas of using probability distributions and deterministic sets and find a unified approach to solve robust portfolio management problems.

Distributionally robust optimization is such a technique that tries to bridge the gap between robust optimization and stochastic programming [34]. It assumes that although the true distribution of random variable is unknown, we can define an uncertainty set for the distribution as we do for the random variable in the classical robust optimization. In other words, we consider a distribution set for the unknown distribution of random variable and try to find the solution that is robust to all feasible distributions in that set.

The primary work of distributionally robust optimization was done by Scarf [62] who suggested a robust formulation for inventory problems based on stochastic programming. In particular, Scarf considered a set of possible distributions of random inventory demand and found the solution with respect to the worst-case of expected cost. In recent years, with the explosive growth in robust optimization, distributionally robust optimization is reemphasized and extensively studied.

Similar to the classical robust optimization, the choice of distribution set is very important in distributionally robust optimization. According to Gabrel et al. [31] and Hanasusanto et al. [40], a large group of distribution sets is based on moment information, which either gives bounds for the moments or is the collection of all distributions with the same first two moments. For detailed information about different kinds of distribution sets, the reader is referred to [26, 40].

Compared to classical robust optimization, generally it is not easy to derive the robust counterparts using distribution sets due to their complex structures [40]. The recent studies focus on finding tractable reformulations or unified frameworks to solve distributionally robust problems. Calafiore and Ghaoui [17] studied a set of radial distributions and showed that distributionally robust problems can be formulated as second order cone problems. Goh and Sim [34] suggested linear and

piece-wise linear decision rules and used distributionally robust bounds to reformulate stochastic linear programs into deterministic convex programs. Wiesemann et al. [65] introduced standardised distribution sets and provided tractable approximations. Hanasusanto et al. [40] proposed a unified framework to solve a standard distributionally robust problem using different distribution sets.

In terms of portfolio management, Popescu [56] studied a robust portfolio management problem to maximize the worst-case expected utility of the portfolio return ($u(\tilde{\mathbf{r}},\mathbf{x})$) and modeled the distribution uncertainty as a set of distributions with fixed first and second moments ($\mathbb{P} \sim (\boldsymbol{\mu},\boldsymbol{\Sigma})$). In order to maintain the consistency of this chapter, here we consider the simplest utility function $u(\tilde{\mathbf{r}},\mathbf{x}) = \tilde{\mathbf{r}}'\mathbf{x}$ that results in the following max-min problem:

$$\max_{\mathbf{x}\in\mathbf{X}} \min_{\mathbb{P}\sim(\boldsymbol{\mu},\boldsymbol{\Sigma})} \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{r}}'\mathbf{x}]. \tag{12.5}$$

Popescu [56] showed that the inner minimization problem in (12.5) is equivalent to an univariate moment problem:

$$\min_{\mathbb{P}\sim(\boldsymbol{\mu}_{\mathbf{x}},\sigma^2_{\mathbf{x}})} \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{r}}'\mathbf{x}],$$

where $\boldsymbol{\mu}_{\mathbf{x}} = \boldsymbol{\mu}'\mathbf{x}$ and $\sigma^2_{\mathbf{x}} = \mathbf{x}'\boldsymbol{\Sigma}\mathbf{x}$. Delage and Ye [24] argued that the assumption of fixed first and second moments would result in sensitive solutions to the noise of data. Alternatively, they considered a general distribution set which takes the estimation error of the first two moments into account:

$$\mathbb{P} \in \mathscr{D}(\mathscr{U}_{\tilde{\mathbf{r}}},\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}},\Upsilon_1,\Upsilon_2) = \left\{ \begin{array}{l} \tilde{\mathbf{r}} \in \mathscr{U}_{\mathbf{r}} \\ (\mathbb{E}[\tilde{\mathbf{r}}] - \hat{\boldsymbol{\mu}})'\hat{\boldsymbol{\Sigma}}^{-1}(\mathbb{E}[\tilde{\mathbf{r}}] - \hat{\boldsymbol{\mu}}) \leq \Upsilon_1 \\ \mathbb{E}[(\tilde{\mathbf{r}} - \hat{\boldsymbol{\mu}})(\tilde{\mathbf{r}} - \hat{\boldsymbol{\mu}})'] \preceq \Upsilon_2\hat{\boldsymbol{\Sigma}} \end{array} \right\}, \tag{12.6}$$

where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are estimated mean and covariance of $\tilde{\mathbf{r}}$, Υ_1 and Υ_2 are some predefined parameters used to adjust the size of the set (can be regarded as the budget of robustness). Notice that the uncertainty set in (12.6) contains all distributions with similar (not required to be exactly the same) first two moments to some degrees Υ_1 and Υ_2. Hence, the inner minimization problem of (12.5) becomes:

$$\begin{array}{lll} \min\limits_{\mathbb{P}\in\mathscr{D}} & & \mathbb{E}[\tilde{\mathbf{r}}'\mathbf{x}] \\ s.t. & (\mathbb{E}[\tilde{\mathbf{r}}] - \hat{\boldsymbol{\mu}})'\hat{\boldsymbol{\Sigma}}^{-1}(\mathbb{E}[\tilde{\mathbf{r}}] - \hat{\boldsymbol{\mu}}) \leq \Upsilon_1 \\ & \mathbb{E}[(\tilde{\mathbf{r}} - \hat{\boldsymbol{\mu}})(\tilde{\mathbf{r}} - \hat{\boldsymbol{\mu}})'] \preceq \Upsilon_2\hat{\boldsymbol{\Sigma}} \\ & \tilde{\mathbf{r}} \in \mathscr{U}_{\tilde{\mathbf{r}}}. \end{array} \tag{12.7}$$

Delage and Ye [24] showed that using the definition of a probability distribution $\mathbb{P}$, (12.7) can be written as a semi-infinite programming problem.

Recently, distributionally robust optimization has also been applied in portfolio optimization with ambiguous utility functions. Since only limited information about investor's risk preference is available when investment decisions are made (e.g. col-

lected from questionnaire), the 'true' utility function of the investor cannot be known exactly [10]. Alternatively, the 'true' utility function is assumed to belong to a set of utility functions (e.g. [42]). Then the investment decision is chosen with respect to the worst-case utility function in this set. In this case, the decision is robust to the ambiguity of the 'true' utility function.

Specifically, Hu et al. [42] suggested an utility function set described by boundary constraints and solved a robust portfolio management problem which maximized the worst-case expected utility within the set. Bertsimas and O'Hair [10] developed a robust and integer optimization approach to learn preferences under data noise and showed that the size of this utility function sets can be reduced when new information comes. Haskell et al. [41] used a distributional robust approach to incorporate ambiguities in both risk preferences and underlying probability distribution of uncertain variables. Armbruster and Delage [2] also introduced a distributional robust approach to find the worst-case utility under incomplete information.

12.7 Robust Risk Measures

The mean-variance portfolio management problem applies a moment-based risk measure that is the variance of portfolio return [54]. One of the main criticisms on the variance risk measure is that it treats both downside and upside movements of asset returns equally. However, the investors may be concerned with the downside risk (capital loss) and the uptrend (capital gain) differently. The empirical studies show that such a risk preference cannot be captured by the variance of portfolio return. Quantile-based risk measures (such as value-at-risk and conditional value-at-risk) consider the probability of loss rather than the gain. It displays some advantages over the moment-based risk measure. In this section, we briefly describe the robust formulations of quantile-based risk metrics.

Robust Value-at-Risk The Value-at-Risk (VaR) measure considers the probability of losses. It is defined as the minimal level ζ such that the probability that the portfolio loss $(-\tilde{\mathbf{r}}'\mathbf{x})$ exceeds ζ is below ε (see for instance, [49]). This can be mathematically formulated as:

$$VaR_{\varepsilon}(\mathbf{x}) = \left\{\min\ \zeta\ \ s.t.\ \ Prob(\zeta \leq -\tilde{\mathbf{r}}'\mathbf{x}) \leq \varepsilon\right\},$$

where the loss probability level is $\varepsilon \in (0, 1]$.

In terms of robust portfolio management, Ghaoui et al. [33] provided a tractable formulation of the worst-case VaR (labeled as WVaR). We now briefly summarize their formulation in this section. Assume that the distribution $\mathbb{P}$ of random asset returns $\tilde{\mathbf{r}}$ is partially known. It belongs to a distribution set $\mathscr{P}$ that contains many possible distributions. Then, the worst-case VaR (WVaR) with respect to the set of probability distributions is defined as:

$$WVaR_{\varepsilon}(\mathbf{x}) = \left\{\min\ \zeta\ \ s.t.\ \ \sup_{\mathbb{P}\in\mathscr{P}} Prob(\zeta \leq -\tilde{\mathbf{r}}'\mathbf{x}) \leq \varepsilon\right\}.$$

The robust portfolio optimization problem that minimizes the worst-case VaR is formulated as:

$$\min_{\mathbf{x}\in\mathbf{X}} \quad WVaR_{\varepsilon}(\mathbf{x}).$$

Ghaoui et al. [33] derived the robust model in view of a set of probability distributions with known first two moments $\hat{\mu}$ and $\hat{\Sigma}$ of asset returns $\tilde{\mathbf{r}}$ using a semidefinite programming problem. This problem was then extended by other researchers by taking into account realistic factors. For instance, Huang et al. [43] considered a WVaR minimization problem with random investment horizon by incorporating the uncertainty in the distribution of exit time into the model. This idea was used later in the conditional value-at-risk context by Huang et al. [44]. Unlike the classical portfolio management problems where the investment horizon is fixed (e.g. 1 month), Huang et al. [44] argued that in reality, the investor can flexibly decide the time to exit the market due to unexpected events in the market or personal reasons. Natarajan et al. [53] developed an asymmetry-robust VaR problem in view of the asymmetric uncertainty set. Goh et al. [35] introduced the so-called partitioned Value-at-Risk measure, which was also motivated by asymmetric property of asset return. Their results showed that the investment portfolio can perform better when the random asset returns are asymmetrically distributed.

Robust Conditional Value-at-Risk Conditional Value-at-Risk (CVaR) is an alternative risk metric to measure a loss distribution and possesses superior mathematical properties comparing to VaR. It is defined as the conditional expectation of all possible losses above VaR [3]. According to Rockafellar and Uryasev [59], CVaR is mathematically formulated as follows:

$$CVaR_{\varepsilon}(\mathbf{x}) = \frac{1}{1-\varepsilon}\int_{-\tilde{\mathbf{r}}'\mathbf{x}\geq VaR_{\varepsilon}(\mathbf{x})} (-\tilde{\mathbf{r}}'\mathbf{x}) f(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}},$$

where $f(\tilde{\mathbf{r}})$ is the probability density function of $\tilde{\mathbf{r}}$.

Pflug [55] proved that CVaR as a coherent risk measure satisfies four axioms: translation invariance, subadditivity, positive homogeneity and monotonicity. For a detailed description of coherent risk measures, the reader is referred to [3]. These four axioms ensure that the CVaR minimization problem is a convex programming problem. Rockafellar and Uryasev [60] showed that the CVaR minimization problem can be formulated as a linear programming problem where each continuous distribution $\mathbb{P}$ is approximated by a discrete distribution. Then a sampling technique can be used to generate a sample space with certain probabilities.

For random asset return $\tilde{\mathbf{r}}$, let's consider J realizations $\mathbf{r}_1, \cdots, \mathbf{r}_J$ with the associated probabilities p_i such that $\sum_{i=1}^{J} p_i = 1$. The CVaR formulation can be approximated by

$$\begin{aligned} CVaR_{\varepsilon}(\mathbf{x}) &= \min_{\zeta}\left\{\zeta + \frac{1}{1-\varepsilon}\mathbb{E}_{\mathbb{P}}\left[\max\{-\tilde{\mathbf{r}}'\mathbf{x}-\zeta, 0\}\right]\right\} \\ &= \min_{\zeta}\left\{\zeta + \frac{1}{1-\varepsilon}\sum_{i=1}^{J} p_i\left[\max\{-\mathbf{r}_i'\mathbf{x}-\zeta, 0\}\right]\right\}. \end{aligned}$$

The worst-case CVaR (WCVaR) can be defined as:

$$WCVaR_{\varepsilon}(\mathbf{x}) = \sup_{\mathbf{p}\in\mathscr{P}} CVaR_{\varepsilon}(\mathbf{x}).$$

Zhu and Fukushima [66] proved that this problem is equivalent to the following min-max problem:

$$WCVaR_{\varepsilon}(\mathbf{x}) = \min_{\zeta} \max_{\mathbf{p}\in\mathscr{P}} \{\zeta + \frac{1}{1-\varepsilon}\sum_{i=1}^{J} p_i[\max\{-\mathbf{r}_i'\mathbf{x}-\zeta, 0\}]\}.$$

By introducing auxiliary variables $s_i = \max\{-\mathbf{r}_i'\mathbf{x}-\zeta, 0\}$ for $i = 1,\ldots,J$, the WCVaR minimization problem can be formulated as a standard linear robust optimization problem:

$$\begin{array}{lll} \min\limits_{\zeta} \max\limits_{\mathbf{p}\in\mathscr{P}} & \zeta + \dfrac{1}{1-\varepsilon}\displaystyle\sum_{i=1}^{J} p_i s_i & \\ s.t. & s_i \geq -\mathbf{r}_i'\mathbf{x} - \zeta, & i = 1,\ldots,J \\ & s_i \geq 0, & i = 1,\ldots,J \\ & \mathbf{e}'\mathbf{x} = 1,\ \mathbf{x} \geq 0. & \end{array}$$

In order to derive the robust CVaR formulation, we need to solve the inner maximization problem in view of an uncertainty set that is constructed by a factor model. For a scaling matrix $\mathbf{G}$ of the probability distribution, let's consider the following uncertainty set

$$\mathscr{P} = \{\mathbf{p} : \mathbf{p} = \mathbf{p}_0 + \mathbf{G}\xi,\ \mathbf{e}'\mathbf{G}\xi = 0,\ \mathbf{p_0} + \mathbf{G}\xi \geq 0,\ \|\xi\| \leq 1\}.$$

The the inner maximization problem becomes

$$\max_{\mathbf{p}\in\mathscr{P}} \zeta + \frac{1}{1-\varepsilon}\mathbf{p}'\mathbf{s} = \zeta + \frac{1}{1-\varepsilon}\left(\mathbf{p}_0'\mathbf{s} + \max_{\mathbf{e}'\mathbf{G}\xi=0,\ \mathbf{p}_0+\mathbf{G}\xi\geq 0,\ \|\xi\|\leq 1} \mathbf{s}'\mathbf{G}\xi\right).$$

Let's assign dual variables ς, σ and τ corresponding to constraints $\mathbf{e}'\mathbf{G}\xi = 0$, $-\mathbf{G}\xi \leq \varpi_0$ and $\|\xi\| \leq 1$, respectively. Using the duality theory, the robust CVaR formulation can be derived as the following deterministic second order conic programming problem:

$$\begin{array}{lll} \min\limits_{\mathbf{x},\zeta,\mathbf{s},\varsigma,\pi,\tau,\upsilon} & \zeta + \dfrac{1}{1-\varepsilon}\left(\mathbf{p}_0'\mathbf{s} - \upsilon + \mathbf{p}_0'\pi\right) & \\ s.t. & s_i \geq -\mathbf{r}_i'\mathbf{x} - \zeta,\ s_i \geq 0, & i = 1,\ldots,N \\ & \mathbf{G}'\mathbf{s} = \mathbf{G}'\mathbf{e}\varsigma + \tau - \mathbf{G}'\pi & \\ & \mathbf{e}'\mathbf{x} = 1,\ \|\tau\| \leq \upsilon & \\ & \mathbf{x}, \pi \geq 0. & \end{array}$$

The reader is referred to Kim et al. [47] for further information on derivation of the robust counterparts of value-at-risk and conditional value-at-risk problems.

Recently, Huang et al. [45] introduced a new risk measure called relative robust CVaR based on relative robust optimization. Unlike the classical robust optimization, the relative robust optimization minimizes the difference between the performance of the solution and relative solutions (e.g. solutions from benchmarks and market competitors).

12.8 Concluding Remarks

In this chapter, we review robust optimization approaches in the context of single-period portfolio allocation problem. We address several issues in robust optimization such as constructing an uncertainty set and derivation of the robust counterpart of the portfolio and risk management problems. We also summarize recent developments in the data-driven and distributional robust optimization approaches and discuss the applicability of these approaches to the single-period portfolio allocation problem. In addition, the robust formulations based on risk metrics such as Value-at-Risk and Conditional Value-at-Risk are presented.

Although the main focus of this chapter is the single stage investment decisions, in practice an asset allocation must be revised at certain time periods and the portfolio needs to be rebalanced over an investment horizon. This requires an extension of the single stage portfolio allocation problem into the multistage and dynamic decision-making framework in which impact of endogenous and/or exogenous uncertain data is taken into account.

The multi-period stochastic programming approaches possess several drawbacks due to existence of inherent inaccuracy and estimation errors as well as computational burdens [38]. It is crucial to develop tractable robust portfolio allocation models in view of more realistic uncertainty sets. Although there exist several applications of robust optimization for the multi-period portfolio allocation problems (see, for instance [7, 11], the data-driven and distributionally robust optimization approaches for the multi-period stochastic programming problems have not been well established. In particular, multi-period robust risk measures for the portfolio selection models are needed. As Chen and Liu [20] has recently stated, there is still gaps in the literature how to ensure the time consistency property when using different risk measures in multi-period problems.

References

1. Anderson, T.W.: An Introduction to Multivariate Statistical Analysis. Wiley, New York (2004)
2. Armbruster, B., Delage, E.: Decision making under uncertainty when preference information is incomplete. Manag. Sci. **61**(1), 111–128 (2015)
3. Artzner, P., Delbaen, F., Eber, J., Heath, D.: Coherent measures of risk. Math. Financ. **9**(3), 203–228 (1999)

4. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. Math. Oper. Res. **23**(4), 769–805 (1998)
5. Ben-Tal, A., Nemirovski, A.: Robust solutions of uncertain linear programs. Oper. Res. Lett. **25**(1), 1–13 (1999)
6. Ben-Tal, A., Nemirovski, A.: Robust solutions of linear programming problems contaminated with uncertain data. Math. Program. **88**(3), 411–424 (2000)
7. Ben-Tal, A., Margalit, T., Nemirovski, A.: Robust modeling of multi-stage portfolio problems. In: High Performance Optimization, pp. 303–328. Kluwer Academic Publisher, Dordrecht, The Netherlands (2000)
8. Ben-Tal, A., Ghaoui, L.E., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
9. Ben-Tal, A., den Hertog, D., Waegenaere, A.D., Melenberg, B., Rennen, G.: Robust solutions of optimization problems affected by uncertain probabilities. Manag. Sci. **59**(2), 341–357 (2012)
10. Bertsimas, D., O'Hair, A.: Learning preferences under noise and loss aversion: an optimization approach. Oper. Res. **61**(5), 1190–1199 (2013)
11. Bertsimas, D., Pachamanova, D.: Robust multiperiod portfolio management in the presence of transaction costs. Comput. Oper. Res. **35**(1), 3–17 (2008)
12. Bertsimas, D., Sim, M.: The price of robustness. Oper. Res. **52**(1), 35–53 (2004)
13. Bertsimas, D., Brown, D.B., Caramanis, C.: Theory and applications of robust optimization. SIAM Rev. **53**(3), 464–501 (2011)
14. Bertsimas, D., Gupta, V., Kallus, N.: Data-driven robust optimization. arXiv preprint arXiv:1401.0212 (2013)
15. Best, M.J., Grauer, R.R.: On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. Rev. Financ. Stud. **4**(2), 315–342 (1991)
16. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer Science & Business Media, New York (2011)
17. Calafiore, G.C., Ghaoui, L.E.: On distributionally robust chance-constrained linear programs. J. Optim. Theory Appl. **130**(1), 1–22 (2006)
18. Campbell, J.Y., Lo, A.W., MacKinlay, A.C.: The Econometrics of Financial Markets. Princeton University Press, Princeton (1997)
19. Ceria, S., Stubbs, R.A.: Incorporating estimation errors into portfolio selection: robust portfolio construction. J. Asset Manag. **7**(2), 109–127 (2006)
20. Chen, Z., Liu, J.: Multi-period robust risk measures and portfolio selection models with regime-switching. Submitted to Oper. Res. (2015)
21. Chen, X., Sim, M., Sun, P.: A robust optimization perspective on stochastic programming. Oper. Res. **55**(6), 1058–1071 (2007)
22. Chopra, V., Ziemba, W.T.: The effects of errors in means, variances, and covariances on optimal portfolio choice. J. Portf. Manag. **19**(2), 6–11 (1993)
23. Dantzig, G.B., Infanger, G.: Multi-stage stochastic linear programs for portfolio optimization. Ann. Oper. Res. **45**(1), 59–76 (1993)
24. Delage, E., Ye, Y.: Distributionally robust optimization under moment uncertainty with application to data-driven problems. Manag. Sci. **58**(3), 595–612 (2010)
25. Duffee, G.: The long-run behavior of firms' stock returns: evidence and interpretations (2002), Working paper
26. Dupacova, J.: Stochastic programming: minimax approach. In: Encyclopedia of Optimization, pp. 3778–3782. Springer, New York (2009)
27. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econ. J. Econ. Soc. **50**(4), 987–1008 (1982)
28. Fama, E.F.: The behavior of stock-market prices. J. Bus. **38**(1), 34–105 (1965)
29. Fama, E.F.: Efficient capital markets: a review of theory and empirical work. J. Financ. **25**(2), 383–417 (1970)
30. Fama, E.F.: Foundations of Finance. Basic Books, New York (1976)

31. Gabrel, V., Murat, C., Thiele, A.: Recent advances in robust optimization: an overview. Eur. J. Oper. Res. **235**(3), 471–483 (2014)
32. Ghaoui, L.E., Lebret, H.: Robust solutions to least-squares problems with uncertain data. SIAM J. Matrix Anal. Appl. **18**(4), 1035–1064 (1997)
33. Ghaoui, L.E., Oks, M., Oustry, F.: Worst-case value-at-risk and robust portfolio optimization: a conic programming approach. Oper. Res. **51**(4), 543–556 (2003)
34. Goh, J., Sim, M.: Distributionally robust optimization and its tractable approximations. Oper. Res. **58**(4), 902–917 (2010)
35. Goh, J., Lim, K.G., Sim, M., Zhang, W.: Portfolio value-at-risk optimization for asymmetrically distributed asset returns. Eur. J. Oper. Res. **221**(2), 396–407 (2012)
36. Goldfarb, D., Iyengar, G.: Robust portfolio selection problems. Math. Oper. Res. **28**(1), 1–38 (2003)
37. Greene, W.H.: Econometric Analysis. Pearson Education India, Delhi (2004)
38. Gulpinar, N., Rustem, B.: Worst-case robust decisions for multi-period mean-variance portfolio optimization. Eur. J. Oper. Res. **183**(3), 981–1000 (2007)
39. Gulpinar, N., Katata, K., Pachamanova, D.A.: Robust portfolio allocation under discrete asset choice constraints. J. Asset. Manag. **12**(1), 67–83 (2011)
40. Hanasusanto, G.A., Roitch, V., Kuhn, D., Wiesemann, W.: A distributionally robust perspective on uncertainty quantification and chance constrained. Math. Program. **151**, 35–62 (2015)
41. Haskell, W.B., Fu, L., Dessouky, M.: Ambiguity in risk preferences in robust stochastic optimization (2015), Working paper
42. Hu, J., Bansal, M., Mehrotra, S.: Robust decision making using a general utility set. Technical Report, Department of IEMS, Northwestern University (2012)
43. Huang, D., Fabozzi, F.J., Fukushima, M.: Robust portfolio selection with uncertain exit time using worst-case VaR strategy. Oper. Res. Lett. **35**(5), 627–635 (2007)
44. Huang, D., Zhu, S.S., Fabozzi, F.J., Fukushima, M.: Portfolio selection with uncertain exit time: a robust cvar approach. J. Optim. Theory Appl. **32**(2), 594–623 (2008)
45. Huang, D., Zhu, S.S., Fabozzi, F.J., Fukushima, M.: Portfolio selection under distributional uncertainty: a relative robust CVaR approach. Eur. J. Oper. Res. **203**(1), 185–194 (2012)
46. Kall, P., Wallace, S.: Stochastic Programming. Wiley, New York (1994)
47. Kim, J.H., Kim, W.C., Fabozzi, F.J.: Recent developments in robust portfolios with a worst-case approach. J. Optim. Theory Appl. **161**(1), 103–121 (2014)
48. Klein, R.W., Bawa, V.S.: The effect of estimation risk on optimal portfolio choice. J. Financ. Econ. **3**(3), 215–231 (1976)
49. Linsmeier, T.J., Pearson, N.D.: Risk measurement: an introduction to value at risk (1996), Working paper
50. Mandelbrot, B.B.: The variation of certain speculative prices. J. Bus. **36**(4), 394–419 (1963)
51. Markowitz, H.: Portfolio selection. J. Financ. **7**(1), 77–91 (1952)
52. Michaud, R.O.: The Markowitz optimization enigma: is 'optimized' optimal? Financ. Anal. J. **45**(1), 31–42 (1989)
53. Natarajan, K., Pachamanova, D., Sim, M.: Incorporating asymmetric distributional information in robust value-at-risk optimization. Manag. Sci. **54**(3), 573–585 (2008)
54. Natarajan, K., Pachamanova, D., Sim, M.: Constructing risk measures from uncertainty sets. Oper. Res. **57**(5), 1129–1141 (2009)
55. Pflug, G.C.: Some remarks on the value-at-risk and the conditional value-at-risk. In: Probabilistic Constrained Optimization, pp. 272–281. Springer, New York (2000)
56. Popescu, I.: Robust mean-covariance solutions for stochastic optimization. Oper. Res. **55**(1), 98–112 (2007)
57. Powell, W.B.: Approximate Dynamic Programming: Solving the Curses of Dimensionality. Wiley Series in Probability and Statistics. Wiley, New York (2011)
58. Prekopa, A.: Stochastic Programming. Springer Science & Business Media, New York (2013)
59. Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. J. Risk **2**, 21–42 (2000)
60. Rockafellar, R.T., Uryasev, S.: Conditional value-at-risk for general loss distributions. J. Bank. Financ. **26**(7), 1443–1471 (2002)

61. Rustem, B., Becker, R.G., Mart, W.: Robust min-max portfolio strategies for rival forecast and risk scenarios. J. Econ. Dyn. Control **24**(11), 1591–1621 (2000)
62. Scarf, H.E.: A min-max solution of an inventory problem. In: Studies in the Mathematical Theory of Inventory and Production, vol. 10, pp. 201–209, Stanford University Press, Stanford (1958)
63. Soyster, A.L.: Convex programming with set-inclusive constraints and application to inexact linear programming. Oper. Res. **21**(5), 1154–1157 (1973)
64. Tutuncu, R.H., Koenig, M.: Robust asset allocation. Ann. Oper. Res. **132**(1–4), 157–187 (2004)
65. Wiesemann, W., Kuhn, D., Sim, M.: Distributionally robust convex optimization. Oper. Res. **62**(6), 1358–1376 (2014)
66. Zhu, S., Fukushima, M.: Worst-case conditional value-at-risk with application to robust portfolio management. Oper. Res. **57**(5), 1155–1168 (2009)
67. Zymler, S., Rustem, B., Kuhn, D.: Robust portfolio optimization with derivative insurance guarantees. Eur. J. Oper. Res. **210**(2), 410–424 (2010)

Chapter 13
Portfolio Optimization with Second-Order Stochastic Dominance Constraints and Portfolios Dominating Indices

Neslihan Fidan Keçeci, Viktor Kuzmenko, and Stan Uryasev

Abstract Portfolio optimization models are usually based on several distribution characteristics, such as mean, variance or Conditional Value-at-Risk (CVaR). For instance, the mean-variance approach uses mean and covariance matrix of return of instruments of a portfolio. However this conventional approach ignores tails of return distribution, which may be quite important for the portfolio evaluation. This chapter considers the portfolio optimization problems with the Stochastic Dominance constraints. As a distribution-free decision rule, Stochastic Dominance takes into account the entire distribution of return rather than some specific characteristic, such as variance. We implemented efficient numerical algorithms for solving the optimization problems with the Second-Order Stochastic Dominance (SSD) constraints and found portfolios of stocks dominating Dow Jones and DAX indices. We also compared portfolio optimization with SSD constraints with the Minimum Variance and Mean-Variance portfolio optimization.

N.F. Keçeci
School of Business, Istanbul University, IU Isletme Fakultesi Avcilar Campus, 34850 Avcilar, Istanbul, Turkey
e-mail: neslihan@istanbul.edu.tr

V. Kuzmenko
Glushkov Institute of Cybernetics, Kyiv, Ukraine
e-mail: kvn_nukr@yahoo.com

S. Uryasev (✉)
Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA
e-mail: uryasev@ufl.edu

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_13

13.1 Introduction

Standard portfolio optimization problems are based on several distribution characteristics, such as mean, variance, and Conditional Value-at-Risk (CVaR) of return distribution. For instance, Markowitz [12] the mean-variance approach uses estimates of mean and covariance matrix of return distribution. Mean-variance portfolio theory works quite well when the return distributions re close to normal.

This chapter considers portfolio selection problem based on the stochastic dominance rule. Stochastic dominance takes into account the entire distribution of return, rather than some specific characteristics. Stochastic dominance produces a partial ordering of portfolio returns and identifies a portfolio dominating some other portfolios [11].

Hadar and Russell [8] demonstrated that a diversified portfolio can dominate a benchmark portfolio in the sense of the Second Order Stochastic Dominance (SSD). Several applications of stochastic dominance theory to portfolio selection are considered by Whitmore and Findlay [16]. Dentcheva and Ruszczynski [3] developed an efficient numerical approach for the portfolio optimization with SSD using partial moment constraints. Roman et al. [14] suggested a portfolio optimization algorithm for SSD efficient portfolios. They used SSD with a multi-objective representation of a problem with CVaR in objective. Kuosmanen [10] and Kopa and Chovanec [9] described SSD portfolio efficiency measure for diversification.

Rudolf and Ruszczynski [15] and Fabian et al. [5, 6] considered cutting plane method to solve optimization problem with SSD constraints. This chapter implemented an algorithm similar to the Rudolf and Ruszczynski [15]. We concentrated on numerical aspects of portfolio optimization with SSD constraints and conducted a case study showing that our algorithm works quite efficiently. We used Portfolio Safeguard (PSG) optimization package of AORDA.com, which has precoded functions for optimization with SSD constraints. We solved optimization problems for stocks in Dow Jones and DAX Indices and found portfolios which SSD dominate these indices. We also compared these portfolios with the Mean-Variance portfolios based on constant and time varying covariance matrices.

13.2 Second Order Stochastic Dominance (SSD)

Let denote by $F_X(t)$ the cumulative distribution function of a random variable X. For two integrable random variables X and Y, we say that X dominates Y in the second-order, if

$$\int_{-\infty}^{\eta} F_X(t)dt \leq \int_{-\infty}^{\eta} F_Y(t)dt, \quad \forall \eta \in \mathbb{R} \tag{13.1}$$

In short we say that X dominates Y in SSD sense and denote it by $X \succcurlyeq_2 Y$ [7]. With the partial moment of a random variable X for a target value η, the SSD dominance is defined as follows

$$E([\eta - X]_+) \leq E([\eta - Y]_+), \quad \forall \eta \in \mathbb{R} \tag{13.2}$$

where, $[\eta - X]_+ = \max(0, \eta - X)$ [13].

13.2.1 SSD Constraints for a Discrete Set of Scenarios

Suppose that Y has a discrete distribution with outcomes, y_i, $i = 1, 2, \ldots, N$. Then the condition (13.2) can be reduced to the finite set of inequalities [13],

$$E([y_i - X]_+) \leq E([y_i - Y]_+), \quad i = 1,, 2, \ldots, N \tag{13.3}$$

We use inequalities (13.3) for obtaining a portfolio X dominating benchmark Y.

13.2.2 Portfolio Optimization Problem with SSD Constraints

Let us denote:

w_j = portfolio weight of the instrument j, $j = 1, \ldots, n$.
p_i = probability of scenario i, $i = 1, \ldots, N$,
r_{ji} = return of instrument j on scenario i,
c_j = cost of investing in instrument $j = 1, \ldots, n$ (estimated return of an instrument is interpreted as negative cost $-c_j$),
$\mathbf{w}$ = vector of portfolio weights, $\mathbf{w} = (w_1, w_2, \ldots, w_n)^\top$,
$\mathbf{r}(\mathbf{w})$ = portfolio return as a function of portfolio weights $\mathbf{w}$,
$c(\mathbf{w})$ = portfolio cost as a function of portfolio weights $\mathbf{w}$.

Portfolio return on scenario i equals:

$$r_i(\mathbf{w}) = \sum_{j=1}^{n} w_j r_{ji}, \quad i = 1, 2, \ldots, N$$

Portfolio cost equals:

$$c(\mathbf{w}) = \sum_{j=1}^{n} c_j w_j$$

The benchmark portfolio Y has a discrete distribution with scenarios y_i, $i = 1, 2, \ldots, N$. We want to find a portfolio SSD dominating the benchmark portfolio Y and having minimum cost $c(\mathbf{w})$. We do not allow for shorting of instruments. Let us denote by W the set of feasible portfolios:

$$W = \{\mathbf{w} \in \mathbb{R}^n : w_j \geq 0, \ j = 1, 2, \ldots, n\}$$

The optimization problem is formulated as follows:

$$\begin{array}{ll} \text{minimize}_{\mathbf{w}} & c(\mathbf{w}) \\ \text{subject to:} & \mathbf{r}(\mathbf{w}) \succeq_2 Y \\ & \mathbf{w} \in W \end{array} \tag{13.4}$$

Since the benchmark portfolio has a discrete distribution, with (13.3) we reduce the portfolio optimization problem (13.4) to:

$$\begin{array}{ll} \text{minimize}_{\mathbf{w}} & \sum_{j=1}^{n} c_j w_j \\ \text{subject to:} & E([y_i - r(\mathbf{w})]_+) \leq E([y_i - Y]_+),\ i = 1,\ldots,N \\ & w_j \geq 0, \qquad\qquad j = 1,\ldots,n \end{array} \tag{13.5}$$

A solution of the optimization problem (13.5) yields a portfolio dominating Y. The number of scenarios (which can be quite large) determines the number of SSD constraints in this optimization problem. Further, we suggest a procedure for elimination redundant constraints in (13.5).

13.3 Algorithm for Portfolio Optimization Problem with SSD Constraints

This section describes cutting plane algorithm for solving problems with SSD constraints in this study. An overview of the cutting-plane methods for SSD problems can be found in [5, 6, 15]. We start the description of the algorithm with the procedure for removing redundant constraints.

13.3.1 Removing Redundant Constraints

Let us consider benchmark scenarios y_{i_1} and y_{i_2} with indices i_1 and i_2 and denote the right hand side values of the constraints for scenarios in problem (13.5) by $C_{i_1} = E([y_{i_1} - Y]_+)$ and $C_{i_2} = E([y_{i_2} - Y]_+)$. If $y_{i_1} \leq y_{i_2}$ and $C_{i_1} \geq C_{i_2}$, for scenarios i_1 and i_2, then constraint i_1 is redundant and it can be removed from the constraint set. This procedure dramatically reduces the number of constraints in the optimization problem (13.5).

13.3.2 Cutting Plane Algorithm

Here are the steps of the algorithm for solving optimization problem (13.5). We denote by s the iteration number.

Step 1. Initialization: $s = 0$. Assign an initial feasible set

$$W_0 = \{\mathbf{w} \in \mathbb{R}^n : w_j \geq 0,\ j = 1,2,\ldots,n\}$$

Assign an initial verification set

$$V_0 = \{\mathbf{w} \in \mathbb{R}^n : E([y_i - r(\mathbf{w})]_+) \leq E([y_i - Y]_+),\ i = 1,2,\ldots,N\}$$

Step 2. Solve the optimization problem

$$\begin{aligned} &\text{minimize}_{\mathbf{w}} \sum_{j=1}^{n} c_j w_j \\ &\text{subject to: } \mathbf{w} \in W_s \end{aligned} \tag{13.6}$$

If all constraints defining the set V_s are satisfied, then the obtained point is optimal to problem (13.5). Otherwise, go to Step 3.
Step 3. Find constraint in V_s with the largest violation and remove it from V_s. Denote this new set of constraints by V_{s+1} (after removing the constraint with the largest violation). Add removed constraint to the constraints defining set W_s and denote it by W_{S+1}. Increase the iteration counter $s = s + 1$ and go back to Step 2.

13.3.3 PSG Code for Optimization with SSD Constraints

The problem (13.5) can be directly solved with the Portfolio Safeguard (PSG) without any additional coding. Here is the code, which can be downloaded from this link.[1]

```
maximize
      avg_g(matrix_sde)
Constraint:= 1
     linear(matrix_budget)
MultiConstraint: <= vector_ubound_sd
      pm_pen (vector_benchmark_sd, matrix_sde)
Box: >= 0, <= 1
```

We have done the case study in PSG MATLAB Environment running many optimizations iteratively. However, here we provided just one code in PSG Run-File format to show that the SSD constrained problems can be easily coded and solved.

[1] Three example problems containing input data and solutions in PSG format are at the following link (see, Problem 1, Dataset 1, 2, 3): http://goo.gl/Fooals.

13.4 Case Study

We solved problem (13.5) with two data sets. The first dataset includes stocks from the Dow Jones (DJ) index and DJ index is considered as a benchmark. Similar, the second dataset includes stocks from the DAX index and the DAX index is used as a benchmark. The data were downloaded from the Yahoo Finance (http://finance.yahoo.com) and include 2500 historical daily returns of stocks from March 24, 2005 to Feb 27, 2015 for DJ index and from April 25, 2005 to Feb 27, 2015 for DAX index. The lists of stocks in indices are taken on March 2, 2015. Therefore, we considered only 29 stocks from the DJ index and 26 stocks from the DAX Index (the appendix contains the list of the stocks selected for this chapter). The stock returns on daily basis (r_{ji}) were calculated using logarithm of ratio of the stock adjusted closing prices (f_i),

$$r_{ji} = \ln(f_i / f_{i-1})$$

We adjusted the stocks prices of four companies from DAX Index.[2] Daily returns are considered as equally probable scenarios in the study.

The optimization problem with SSD constraint (in this case study) finds a portfolio SSD dominating the benchmark and having maximum expected portfolio return. Shorting is not allowed. The sum of portfolio weights is equal to 1,

$$\sum_{j=1}^{n} w_j = 1, \quad w_j \geq 0, \quad j = 1, 2, \ldots, n$$

We compared performance of the SSD based portfolios with Equally Weighted, Minimum Variance and Mean-Variance portfolios with the constant and time-varying covariance matrices. Here is a brief description of portfolios:

1. Equally Weighted (EW): All stocks in the portfolio are equally weighted. Every stock has same weight $(1/n)$, where n is the number of stocks in the portfolio.
2. Minimum Variance (MinVar): Minimum Variance portfolio has minimum variance without any constraint on portfolio return. Shorting is not allowed and the sum of the portfolio weights is equal to 1.
3. Mean-Variance (Mean-Var): Mean-variance portfolio [12] uses mean return and the variance of the stock returns. The approach finds efficient portfolios having minimum variance for a desired level of portfolio return or equivalently having maximum portfolio return for a given variance. We considered Mean-Var problems having variance in the objective function and the expected portfolio return 12 % per year in the constraint, and 0.2 upper bound constraint on the positions. Shorting is not allowed and the sum of the portfolio weights is equal to 1.

The classical Mean-Variance model considers the constant covariance matrix. We also considered the time dependent covariance matrix using DCC-GARCH model in MinVar and Mean-Var approaches. Further, we briefly describe the estimation procedure for the time-dependent covariance matrix.

[2] DB1.DE, FRE.DE, IFX.DE and MRK.DE stock prices are adjusted for splits.

13.4.1 Estimation of Time-Varying Covariance Matrix

We considered a dynamic conditional correlation DCC-GARCH (DCC) model for the estimation of large time-dependent covariance matrices [4]. We estimated the time dependent covariance matrix using DCC-GARCH model (assuming that correlations may change over time). The time-dependent covariance matrix H_t is extracted from the DCC-GARCH model, where $H_t = D_t R_t D_t$. Here, D_t is the diagonal matrix from a univariate GARCH model and R_t is the time dependent correlation matrix. This chapter assumes the simplest conditional mean return equation where $\overline{r}_j = N^{-1} \sum_{i=1}^{N} r_{ji}$ is the sample mean and the deviation of returns $(r_t - \overline{r})$ is conditionally normal with zero mean and time-dependent covariance matrix H_t [2]. We consider the time-dependent covariance matrix H_t in a simple DCC(1,1)-GARCH model. We used H_t in MinVar and Mean-Var problems.

The next Sect. 13.4.2 compares SSD constrained optimization with the MinVar and the Mean-Var approaches for all available historical data in a static setting. The code was implemented with MATLAB R2012b. We have used PSG riskprog function in MATLAB environment to solve MinVar and Mean-Variance portfolio problems. For the estimation of the time-dependent covariance matrix we have used MFE Toolbox[3]. The Sect. 13.4.3 we compares out-of-sample performance of portfolios in time series framework. The calculations were performed on a computer having 3.4 GHz CPU and 8 GB of RAM.

13.4.2 Comparing Numerical Performance of Various Portfolio Settings

We benchmarked the cutting plane algorithm described in Sect. 13.3.2 with the direct PSG code described in Sect. 13.3.3. We got the same results with both approaches. Further in tables we report performance of the direct PSG code. The dataset includes 2500 historical daily stock returns. Firstly we optimized portfolios with all approaches using available 2500 historical daily returns.

Table 13.1 shows the expected yearly returns of portfolios for all considered approaches.

The SSD dominating portfolios can be used for actual investments. At least in the past, these portfolios SSD dominated the corresponding indices. Moreover, the expected yearly return of the portfolio SSD dominating the DJ index equals 0.10029 and significantly exceeds the DJ index return in this period. Similar observations are valid for the portfolio of DAX index; the expected yearly return of portfolio SSD dominating the benchmark equals 0.14894.

We compared solving times of SSD constraint optimization problem (using direct PSG optimization) with the MinVar and Mean-Var approaches (using PSG *riskprog*

[3] DCC-GARCH models are estimated with Kevin Sheppard's (Multivariate GARCH) MFE Toolbox. http://www.kevinsheppard.com/MFE_Toolbox

Table 13.1: Expected yearly returns of portfolios

Portfolios	DJI	DAX
EW	0.09617	0.10179
MinVar	0.08668	0.14135
Mean-Var	0.12693	0.12693
DCC MinVar	0.09034	0.13466
DCC Mean-Var	0.12693	0.12693
SSD	0.10029	0.14894
Benchmark	0.05682	0.10484

subroutine). Data loading and solving times are given in Table 13.2. The optimization is done almost instantaneously and data loading takes some fraction of a second. The time-dependent covariance matrix estimation with MFE Toolbox additionally takes about 30 s (for MinVar and Mean-Var optimization).

Table 13.2: Loading and solving times (in seconds) with PSG in MATLAB Environment

	DJ		DAX	
Problem	Loading	Solving	Loading	Solving
SSD constrained (PSG code)	0.24	0.01	0.23	0.01
MinVar (PSG riskprog)	0.22	0.01	0.23	0.01
Mean-Var (PSG riskprog)	0.31	0.01	0.32	0.01

13.4.3 Out-of-Sample Simulation

Secondly, we have evaluated the out-of-sample performance of considered approaches. We considered a time series framework where the estimation period (750 and 1000 days) is rolled over time. Portfolios are re-optimized on every first business day of the month using the recent historical daily returns (750 or 1000). We kept constant positions during the month. Regarding the return constraint in the Mean-Var problem, if the expected return 12 % per year is not feasible (in the beginning of the month), than we set 6 % expected return constraint and if we still do not have feasibility, we reduce the expected return to 3 %, and then to 0 %. A difficulty in estimation of the covariance matrices with DCC model is that the time-dependent conditional correlation matrix has to be positive definite for all time moments [1]. We observe that with a small in-sample time intervals (such as 250 days) the variance-covariance matrix may not be positive-definite. Therefore, we have used 750 and 1000 days in-sample periods.

Table 13.3 shows out-of-sample total compounded returns of considered portfolios. In particular, we observe that the SSD constrained portfolio for DJ stocks with 750 days in-sample and DAX stocks with 1000 days in-sample, have highest compounded returns among all portfolios.

Table 13.3: Out-of-sample total compounded returns of the portfolios

	DJ		DAX	
Portfolios	750	1000	750	1000
EW	1.6698	2.7560	1.2120	1.9022
MinVar	1.4787	2.0035	1.6186	2.4410
Mean-Var	1.6715	2.1487	1.4912	2.6219
DCC MinVar	1.3876	2.0463	1.7763	2.2062
DCC Mean-Var	1.6499	2.2556	1.4344	2.3435
SSD	1.8817	2.1692	1.3987	2.8827
Benchmark	1.2729	2.2275	1.3399	1.8674

Figures 13.1, 13.2, 13.3, and 13.4 show the out-of-sample compounded daily returns of the portfolios.

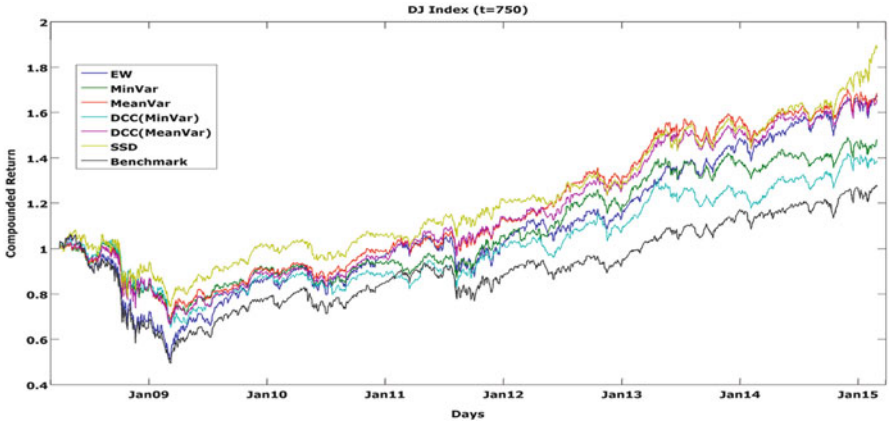


Fig. 13.1: Compounded (on daily basis) returns of portfolios including DJ stocks, $t = 750$

The out-of-sample performances of portfolios are represented in Tables 13.4, 13.5, 13.6, and 13.7. The tables include yearly compounded portfolio returns for 2009–2014 years, the total compounded portfolio return (T_R) and Sharpe Ratio (Sh_R).

- Table 13.4 (DJ stocks, $t = 750$). SSD constrained portfolio has the highest Total Compounded Return (1.8817) and Sharpe Ratio (0.7906).
- Table 13.5 (DJ stocks, $t = 1000$). SSD constrained portfolio has Sharpe Ratio (1.2771) higher than the all considered portfolios except Equally Weighted portfolio.

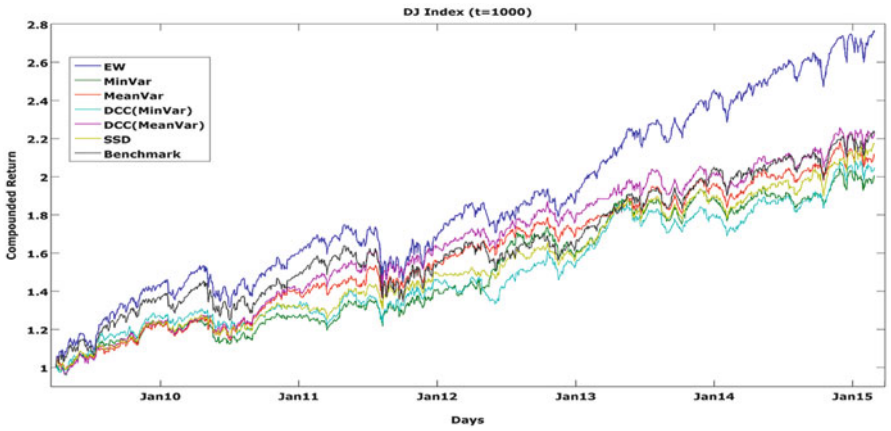


Fig. 13.2: Compounded (on daily basis) returns of portfolios including DJ stocks, $t = 1000$

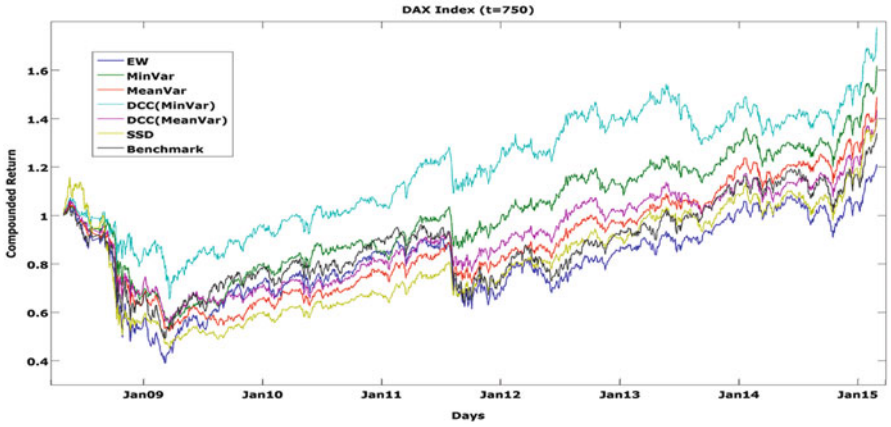


Fig. 13.3: Compounded (on daily basis) returns of portfolios including DAX stocks, $t = 750$

- Table 13.6 (DAX stocks, $t = 750$). SSD constrained portfolio has Sharpe Ratio (0.2732) and Total Compounded Return (1.3987) higher than the Benchmark and Equally Weighted portfolios.
- Table 13.7 (DAX stocks, $t = 1000$). SSD constrained portfolio has the highest Total Return (2.8827) and Sharpe Ratio (1.3523).

Table 13.8 shows weights of SSD constrained portfolios at the last month of the out-of-sample period. Also, the table shows SSD dominating portfolios over all in-sample 2500 days. The table shows only stocks with non-zero positions.

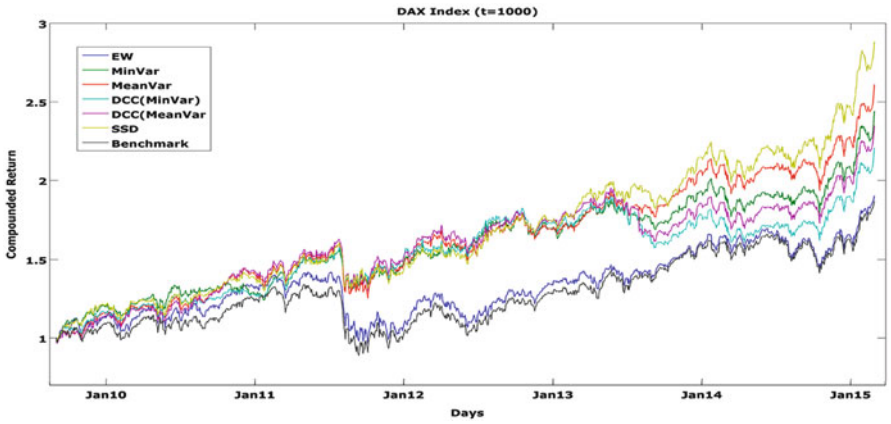


Fig. 13.4: Compounded (on daily basis) returns of portfolios including DAX stocks, $t = 1000$

Table 13.4: Yearly compounded returns, total compounded return (T_R), Sharpe ratio (Sh_R) for DJ stocks ($t = 750$)

Portfolios	2009	2010	2011	2012	2013	2014	T_R	Sh_R
EW	1.2257	1.1256	1.0467	1.1249	1.3063	1.1029	1.6698	0.5533
MinVar	1.0903	1.0282	1.1259	1.1158	1.1814	1.0451	1.4787	0.4584
Mean-Var	1.0618	1.0993	1.1424	1.1399	1.2220	1.0473	1.6715	0.7034
DCC MinVar	1.0277	1.0119	1.1513	1.0612	1.1574	1.1090	1.3876	0.4025
DCC Mean-Var	1.0405	1.0865	1.1864	1.1041	1.2104	1.0678	1.6499	0.6835
SSD	1.0931	1.0493	1.1495	1.0443	1.2089	1.1498	1.8817	0.7906
Benchmark	1.154	1.0958	1.0321	1.0652	1.2584	1.0688	1.2729	0.2239

Table 13.5: Yearly compounded returns, total compounded return (T_R), Sharpe ratio (Sh_R) for DJ stocks ($t = 1000$)

Portfolios	2010	2011	2012	2013	2014	T_R	Sh_R
EW	1.1256	1.0467	1.1249	1.3063	1.1029	2.7560	1.2902
MinVar	1.0294	1.1109	1.1349	1.1852	1.0593	2.0035	0.9230
Mean-Var	1.1241	1.1129	1.0920	1.1993	1.0736	2.1487	1.1965
DCC MinVar	1.0437	1.1064	1.0530	1.1702	1.1321	2.0463	0.9803
DCC Mean-Var	1.1207	1.1548	1.0866	1.1822	1.0899	2.2556	1.2653
SSD	1.0721	1.1327	1.0525	1.2078	1.1153	2.1692	1.2771
Benchmark	1.0958	1.0321	1.0652	1.2584	1.0688	2.2275	1.0165

Table 13.6: Yearly compounded returns, total compounded return (T_R), Sharpe ratio (Sh_R) for DAX stocks ($t = 750$)

Portfolios	2009	2010	2011	2012	2013	2014	T_R	Sh_R
EW	1.2890	1.1677	0.8119	1.2526	1.2034	1.0217	1.2120	0.1245
MinVar	1.1609	1.0742	1.1485	1.1590	1.1537	1.0698	1.6186	0.5457
Mean-Var	0.9763	1.1691	1.0700	1.1896	1.2448	1.0740	1.4912	0.4635
DCC MinVar	1.1231	1.0966	1.1853	1.1346	1.0468	1.0436	1.7763	0.6348
DCC Mean-Var	0.9799	1.1565	1.0793	1.1612	1.1282	1.0789	1.4344	0.3939
SSD	1.0284	1.1304	1.0950	1.2322	1.2031	1.0885	1.3987	0.2732
Benchmark	1.1891	1.1411	0.8181	1.2671	1.2414	1.0123	1.3399	0.2279

Table 13.7: Yearly compounded returns, total compounded return (T_R), Sharpe ratio (Sh_R) for DAX stocks ($t = 100$)

Portfolios	2010	2011	2012	2013	2014	T_R	Sh_R
EW	1.1677	0.8119	1.2526	1.2034	1.0217	1.9022	0.6784
MinVar	1.0714	1.1479	1.1541	1.1371	1.0746	2.4410	1.1886
Mean-Var	1.2238	1.0407	1.1562	1.2211	1.0962	2.6219	1.3164
DCC MinVar	1.1033	1.1825	1.1501	1.0282	1.0606	2.2062	1.0335
DCC Mean-Var	1.2547	1.0542	1.1593	1.0626	1.0867	2.3435	1.0169
SSD	1.1452	1.0343	1.2235	1.2323	1.1425	2.8827	1.3523
Benchmark	1.1411	0.8181	1.2671	1.2414	1.0123	1.8674	0.6676

Table 13.8: SSD constrained portfolios (table shows only selected stocks)

DJ	Weights			DAX	Weights		
Code	750 days	1000 days	2500 days	Code	750 days	1000 days	2500 days
BA	0.06984	0.00808	0.00243	ADS	0.00000	0.00000	0.01287
DIS	0.34832	0.10397	0.00000	BAYN	0.00000	0.00000	0.02686
HD	0.19280	0.19317	0.00000	BEI	0.00000	0.02238	0.20000
IBM	0.00000	0.00000	0.01959	CON	0.33042	0.04653	0.00000
JNJ	0.00000	0.22105	0.20000	DPW	0.06556	0.12844	0.00000
KO	0.00000	0.00000	0.16569	DTE	0.02086	0.00000	0.00000
MCD	0.00000	0.00000	0.20000	FME	0.00000	0.00000	0.20000
MMM	0.08726	0.00000	0.00000	FRE	0.31189	0.43899	0.12149
MRK	0.07009	0.15284	0.00000	LIN	0.00000	0.00000	0.10719
NKE	0.05919	0.05844	0.01776	MRK	0.27127	0.36366	0.12265
PG	0.00000	0.02058	0.19368	MUV2	0.00000	0.00000	0.08076
T	0.00000	0.00000	0.01456	SAP	0.00000	0.00000	0.12819
TRV	0.08155	0.00000	0.00000				
UNH	0.09096	0.08374	0.00000				
VZ	0.00000	0.00000	0.05402				
WMT	0.00000	0.15814	0.13226				

13.5 Conclusions

In this chapter we tested algorithms for portfolio optimization with SSD constraints. The algorithms are very efficient and solve optimization problems nearly instantaneously (solution times less than 0.01 s for the considered cases).

We have done out-of-sample simulations and compared SSD constrained portfolios with the minimum variance and mean-variance portfolios. The portfolios were constructed from the stocks of DJ and DAX indices. SSD constrained portfolio demonstrated quite good out-of-sample performance and in some cases had highest compounded return and Sharpe ratio (among all considered portfolios).

Appendix: Company Codes and Names

	DAX code	Name	DJ code	Name
1	ADS	Adidas AG	AXP	American Express Company
2	ALV	Allianz SE	BA	The Boeing Company
3	BAS	BASF SE	CAT	Caterpillar Inc.
4	BAYN	Bayer AG	CSCO	Cisco Systems, Inc.
5	BEI	Beiersdorf AG	CVX	Chevron Corporation
6	BMW	Bayerische Mot. Werke Aktienges.	DD	E. I. du Pont de Nemours and Company
7	CBK	Commerzbank AG	DIS	The Walt Disney Company
8	CON	Continental Aktiengesellschaft	GE	General Electric Company
9	DAI	Daimler AG	GS	The Goldman Sachs Group, Inc.
10	DB1	Deutsche Boerse AG	HD	The Home Depot, Inc.
11	DBK	Deutsche Bank AG	IBM	Int. Business Machines Corporation
12	DPW	Deutsche Post AG	INTC	Intel Corporation
13	DTE	Deutsche Telekom AG	JNJ	Johnson & Johnson
14	EOAN	E.ON SE	JPM	JPMorgan Chase & Co.
15	FME	Fres. Med. Care AG & Co. KGAA	KO	The Coca-Cola Company
16	FRE	Fresenius SE & Co KGaA	MCD	McDonald's Corp.
17	HEI	HeidelbergCement AG	MMM	3M Company
18	IFX	Infineon Technologies AG	MRK	Merck & Co. Inc.
19	LHA	Deutsche Luft. Aktiengesellschaft	MSFT	Microsoft Corporation
20	LIN	Linde Aktiengesellschaft	NKE	Nike, Inc.
21	MRK	Merck KGaA	PFE	Pfizer Inc.
22	MUV2	Münchener R. G. A.	PG	The Procter & Gamble Company
23	SAP	SAP SE	T	AT&T, Inc.
24	SDF	K+S Aktiengesellschaft	TRV	The Travelers Companies, Inc.
25	SIE	Siemens Aktiengesellschaft	UNH	UnitedHealth Group Incorporated
26	TKA	ThyssenKrupp AG	UTX	United Technologies Corporation
27			VZ	Verizon Communications Inc.
28			WMT	Wal-Mart Stores Inc.
29			XOM	Exxon Mobil Corporation

References

1. Bauwens, L., Laurent, S., Rombouts, J.V.: Multivariate GARCH models: a survey. J. Appl. Econ. **21**(1), 79–109 (2006)
2. Carol, A.: Market Risk Analysis: Practical Financial Econometrics, Hoboken. Wiley, New Jersey (2008)
3. Dentcheva, D., Ruszczyński, A.: Portfolio optimization with stochastic dominance constraints. J. Bank. Financ. **30**(2), 433–451 (2006)
4. Engle, R.F., Sheppard, K.: Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Technical Report, National Bureau of Economic Research (2001)
5. Fábián, C.I., Mitra, G., Roman, D.: Processing second-order stochastic dominance models using cutting-plane representations. Math. Program. **130**(1), 33–57 (2011)
6. Fábián, C.I., Mitra, G., Roman, D., Zverovich, V.: An enhanced model for portfolio choice with SSD criteria: a constructive approach. Quant. Financ. **11**(10), 1525–1534 (2011)
7. Hadar, J., Russell, W.R.: Rules for ordering uncertain prospects. Am. Econ. Rev. **59**(1), 25–34 (1969)
8. Hadar, J., Russell, W.R.: Stochastic dominance and diversification. J. Econ. Theory **3**(3), 288–305 (1971)
9. Kopa, M., Chovanec, P.: A second-order stochastic dominance portfolio efficiency measure. Kybernetika **44**(2), 243–258 (2008)
10. Kuosmanen, T.: Efficient diversification according to stochastic dominance criteria. Manag. Sci. **50**(10), 1390–1406 (2004)
11. Levy, H.: Stochastic Dominance: Investment Decision Making under Uncertainty. Springer, Berlin (2006)
12. Markowitz, H.: Portfolio selection. J. Financ. **7**(1), 77–91 (1952)
13. Ogryczak, W., Ruszczyński, A.: From stochastic dominance to mean-risk models: semideviations as risk measures. Eur. J. Oper. Res. **116**(1), 33–50 (1999)
14. Roman, D., Darby-Dowman, K., Mitra, G.: Portfolio construction based on stochastic dominance and target return distributions. Math. Program. **108**(2–3), 541–569 (2006)
15. Rudolf, G., Ruszczynski, A.: Optimization problems with second order stochastic dominance constraints: duality, compact formulations, and cut generation methods. SIAM J. Optim. **19**(3), 1326–1343 (2008)
16. Whitmore, G.A., Findlay, M.C.: Stochastic Dominance: An Approach to Decision-Making Under Risk. Lexington Books, Toronto (1978)

Chapter 14
Robust DEA Approaches to Performance Evaluation of Olive Oil Production Under Uncertainty

Kazım Barış Atıcı and Nalân Gülpınar

Abstract In this chapter, we are concerned with performance evaluation of olive oil production using Data Envelopment Analysis (DEA) under uncertainty. In order to measure production efficiency of olive-growing farms, we first apply an imprecise DEA approach by taking into account optimistic and pessimistic perspectives on uncertainty realized in olive oil production yield. We then consider robust optimization based DEA under an uncertainty set where the random data belong. The robust DEA model enables to adjust level of conservatism that is defined by the price of robustness of the uncertainty set. The performance of imprecise and robust DEA models is illustrated via a case study of olive-growing farms located in the Aegean Region of Turkey. The numerical experiments reveal that the efficiency scores and efficiency discriminations dramatically depend on how the uncertainty is treated both in imprecise and robust DEA modeling. There exists a trade-off between the protection level and conservatism of the efficiency scores.

14.1 Introduction

Data Envelopment Analysis (DEA) is a well-established non-parametric approach for identifying relative efficiency of organizations or organizational units that are producing multiple outputs through the use of multiple inputs [13, 23]. The DEA approach has been widely applied for the performance evaluation of different

K.B. Atıcı
Department of Business Administration, Hacettepe University, Ankara, Turkey
e-mail: kba@hacettepe.edu.tr

N. Gülpınar (✉)
Warwick Business School, The University of Warwick, Coventry, UK
e-mail: Nalan.Gulpinar@wbs.ac.uk

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8_14

aspects of business practices in various industries. Together with banking, healthcare, transportation and education, agriculture is one of the top-five industries that DEA has been applied [18]. In particular, DEA and related methodologies have been used for identifying relative technical efficiency of various types of agricultural establishments (for recent examples see [1, 2, 9, 15]).

The standard DEA methodology requires perfect information about data. In other words, multiple input and output parameters for each decision making unit are assumed to be known exactly. However, in various real world applications, some or all parameters involve uncertainty. Often little is known about the specific distributions of future uncertainties, and little data are available for estimating the probability distributions of these uncertainties. In many cases, it may be preferable to provide general information about the uncertainties, such as means, ranges, and directional deviations, rather than generating specific scenarios. In this case, they may be represented in the forms of ordinal or bound data.

The standard (deterministic) DEA approach was extended to Imprecise Data Envelopment Analysis (IDEA) to handle data uncertainty by Cooper et al. [11, 12]. The production efficiency of decision-making units is determined in view of such uncertain parameters that are assumed to take either optimistic or pessimistic perspectives. The reader is referred to Zhu [24, 25] and Park [19] for various applications of IDEA.

As we will discuss in more detail later, robust optimization is a technique for decision making under uncertainty that is concerned with finding the optimal solution when uncertain parameters in the problem take their worst-case values in prespecified uncertainty sets. Robust optimization was independently developed by Ben-Tal and Nemirovski [4] and Ghaoui and Lebret [16], and has experienced tremendous growth in the last decade mainly because of computational tractability and practical implementation (for example, see [5, 7, 8]).

Robust optimization has also been adopted to DEA for handling data uncertainty arising in input and output parameters. However, the robust DEA framework has not been yet widely applied in practice. Sadjadi and Omrani [20] considered robustifying uncertainty on output parameters for the performance assessment of electricity distribution companies. Shokouhi et al. [21] proposed a tractable robust approach for imprecise DEA where both input and output parameters are constrained within an uncertainty set. They applied a Monte-Carlo simulation to illustrate performance of the robust DEA model using a small example.

This chapter focuses on an agricultural performance evaluation problem under uncertainty using DEA. More precisely, the imprecise and robust DEA models are developed to measure technical efficiency of olive growing farms. To the best of our knowledge, this research is the first attempt to model olive oil production problem under uncertainty using the IDEA and robust DEA approaches. It is worthwhile to mention that imprecise DEA is applied to the olive oil production problem rather than standard (deterministic) DEA due to stochastic nature of output parameters associated with farms' olive oil yields. Uncertainty is represented as in the form of bound data varying dependently on the olive production. We apply both optimistic [11] and pessimistic [19] perspectives for data uncertainty within the imprecise DEA framework.

We also apply the robust DEA approach to find the worst-case production efficiency of olive-growing farms in view of uncertain olive oil production level varying within a pre-specified uncertainty set. A tractable robust DEA model is derived using an uncertainty set, introduced by Bertsimas and Sim [7], for random output parameters. This model enables to adjust the level of conservatism that is defined by the price of robustness of the uncertainty set.

We consider a real world case study of olive oil producing farms located in the Aegean Region of Turkey in order to illustrate performance of imprecise and robust DEA models. The production performance of those farms is measured in terms of efficiency scores under data uncertainty. We study how efficiency scores change between imprecise and robust DEA modeling. In addition to the comparison of efficiency scores in robust and imprecise DEA models, we also investigate the impact of the size of uncertainty sets and model parameters on the robust and imprecise DEA scores of smaller groups of farms via simulating the estimated olive oil production and its deviations.

The rest of this chapter is organized as follows. Section 14.2 provides an insight on imprecise DEA models. In Sect. 14.3, we present a brief introduction to robust optimization modeling of DEA and derive mathematical formulations of robust DEA models. Section 14.4 focuses on the case study and describes the data set in terms of input and output variables. Section 14.5 presents the numerical results of relative production efficiency obtained through imprecise and robust DEA approaches. Finally, Sect. 14.6 summarizes our findings.

14.2 DEA Modeling

This section is a brief introduction to deterministic and imprecise DEA modeling. We consider imprecise DEA as a benchmark approach for the olive oil production problem under uncertainty. Before formulating the imprecise DEA model, let us describe a standard DEA linear program since it is a fundamental model for both imprecise and robust DEA approaches.

14.2.1 Deterministic DEA Model

As mentioned before, DEA is used to measure relative efficiency of a decision making unit with respect to other units producing multiple outputs through the use of multiple inputs. The fundamental model (referred to the CCR DEA model) was introduced by the original work of Charnes et al. [10]. The CCR DEA model basically builds on the idea of maximizing the ratio of weighted combination of outputs to weighted combination of inputs.

Let us consider N decision making units. We assume that each decision making unit j (for $j = 1, \cdots, N$) uses M different inputs x_{ij} (for $i = 1, \cdots, M$) and produces S different outputs y_{rj} (for $r = 1, \cdots, S$). Let μ_r and w_i denote the weights (or decision

variables) corresponding to output r and input i, respectively. The CCR DEA model calculates the efficiency score for the decision making unit o under consideration by solving the following linear problem:

$$\begin{aligned} \max \ & \sum_{r=1}^{S} \mu_r y_{ro} \\ \text{s.t.} \ & \sum_{i=1}^{M} w_i x_{io} = 1 \\ & \sum_{r=1}^{S} \mu_r y_{rj} - \sum_{i=1}^{M} w_i x_{ij} \leq 0, \ j = 1, \cdots, N \\ & \mu_r, w_i \geq 0, \qquad\qquad i = 1, \cdots, M, \ \ r = 1, \cdots, S. \end{aligned} \tag{14.1}$$

Notice that the standard DEA model (14.1) is constructed by using exact (deterministic) data values for input and output parameters for each decision making unit. However, it is not always possible to have perfect information about the data related to input and/or output values of decision making units. The gathered data may involve inaccuracy due to estimation error, and even data uncertainty may exist due to the nature of the underlying problem.

For these cases, Cooper et al. [11] first introduced the concept of imprecise data into the DEA framework. The term "imprecise data" reflects the situation where some of the input and output data are only known to lie within bounded intervals [14]. Thus, Imprecise Data Envelopment Analysis (IDEA) permits the incorporation of bounded or ranked data into the DEA models.

14.2.2 Imprecise DEA Model

Let D_r^+ and D_i^- denote sets for the input and output parameters including both imprecise and exact data. The values of y_r and/or x_i are not known exactly, but need to be determined in sets D_r^+ and D_i^-. Then the IDEA model based on the CCR DEA formulation is stated as follows.

$$\begin{aligned} \max \ & \sum_{r=1}^{S} \mu_r y_{ro} \\ \text{s.t.} \ & \sum_{i=1}^{M} w_i x_{io} = 1 \\ & \sum_{r=1}^{S} \mu_r y_{rj} - \sum_{i=1}^{M} w_i x_{ij} \leq 0, \ j = 1, \cdots, N \\ & y_r = (y_{rj}) \in D_r^+ \qquad r = 1, \cdots, S \\ & x_i = (x_{ij}) \in D_i^- \qquad i = 1, \cdots, M \\ & \mu_r, w_i \geq 0, \qquad\qquad i = 1, \cdots, M, \ \ r = 1, \cdots, S, \end{aligned} \tag{14.2}$$

where the decision variables μ_r and w_i represent weights corresponding to output and input data as defined before.

For the deterministic DEA model, the exact values of y_r and/or x_i taken from sets D_r^+ and D_i^- are simply substituted in (14.2). This leads to a linear program as stated by the standard DEA model (14.1). However, the IDEA model (14.2) becomes a nonlinear optimization problem. In order to solve the nonlinear programming problem, Cooper et al. [11] and Kim et al. [17] converted the nonlinear model into a linear program via scale transformations and variable alterations. Cooper et al. [12] then applied the IDEA model (where an imprecise output parameter for all units is defined in the form of intervals) to measure performance efficiency of a mobile telecommunication company in Korea. Zhu [25] also considered the IDEA model for the same telecommunication problem, but solved it as a standard DEA problem. He showed that the same efficiency scores as in Cooper et al. [12] can be obtained by simply substituting the output parameter of the unit under evaluation to its upper bound while fixing the output parameters of the remaining units to their lower bounds of the corresponding intervals. Therefore, the unit under evaluation is assumed to perform at its best (as fixed at the upper bound of the corresponding interval of the output parameters) while the other units are assumed to perform at the worst-case (as fixed at the lower bounds of the interval of output parameters).

Park [19] proved that the IDEA formulation (introduced by Cooper et al. [11]) in fact produces an "*optimistic*" strategy since the efficient score is evaluated at the best scenario (selected within pre-specified imprecise data interval) available for the underlying decision making unit. Therefore, the objective function for the IDEA model to achieve the optimistic strategy can be formulated as follows:

$$\max_{y_r \in D_r^+, x_i \in D_i^-} \max_{\mu, w} \sum_{r=1}^{S} \mu_r y_{ro}$$

Similarly, a "*pessimistic*" strategy within the IDEA context is achieved by the following min-max objective function

$$\min_{y_r \in D_r^+, x_i \in D_i^-} \max_{\mu, w} \sum_{r=1}^{S} \mu_r y_{ro}$$

The IDEA model with the min-max objective transforms the bounded data to exact data so that the model seeks to evaluate the unit under evaluation in the worst scenario possible. In other words, the unit under consideration is evaluated by the worst scenario possible (specified at the lower bound of the interval of the imprecise parameter) while the other units perform at their best scenario (specified at the upper bound of the interval of the imprecise parameter). Therefore, the solution of the-min-max optimization problem provides a conservative (worst-case) efficient score for the underlying unit.

If any unit is determined as efficient by both optimistic and pessimistic perspectives within IDEA approach, then it is declared as *"perfectly efficient"*. On the other hand, it is called *"potentially efficient"* when it is efficient under the optimistic strategy and inefficient under the pessimistic strategy [19].

Following the transformations introduced by Soyster [22], the IDEA model in view of the pessimistic perspective is formulated in a general form as follows;

$$
\begin{aligned}
&\max \textstyle\sum_{r=1}^{S} \mu_r \, inf\{y_{ro} \mid y_r \in D_r^+\} \\
&s.t. \quad \sum_{i=1}^{M} w_i \, sup\left\{x_{io} \mid x_i \in D_i^-\right\} = 1 \\
&\qquad \sum_{r=1}^{S} \mu_r \, inf\left\{y_{ro} \mid y_r \in D_r^+\right\} - \sum_{i=1}^{M} w_i \, sup\left\{x_{io} \mid x_i \in D_i^-\right\} \le 0, \qquad (14.3) \\
&\qquad \sum_{r=1}^{S} \mu_r \, inf\left\{y_{rj} \mid y_r \in D_r^+\right\} - \sum_{i=1}^{M} w_i \, sup\left\{x_{ij} \mid x_i \in D_i^-\right\} \le 0, \; j = 1, \cdots, N, \; j \ne o \\
&\qquad \mu_r, w_i \ge 0, \qquad\qquad i = 1, \cdots, M, \;\; r = 1, \cdots, S
\end{aligned}
$$

where the 'sup' and 'inf' are replaced by 'max' and 'min', respectively, when D_r^+ and D_i^- are closed and bounded sets. The reader is referred to Park [19] for further information on the generalized linear program.

14.3 Robust DEA Approach

As mentioned in the introduction, the robust optimization approach to solving an optimization problem with uncertain data involves specifying appropriate uncertainty sets for the uncertain coefficients, and finding a solution that guarantees feasibility even if the uncertain coefficients take their worst-case values within the uncertainty sets. A brief introduction to the main ideas of robust linear optimization (the type of problem with which we are dealing in this chapter) is provided next; for further information, the reader is referred to Ben-Tal and Nemirovski [4, 5] as well as Ben-Tal et al. [6]. We then derive the robust DEA model that is to be applied for the case study of olive oil production problem described in Sect. 14.4.

14.3.1 Robust Linear Optimization

Consider, for example, a linear program

$$\max \left\{ \mathbf{c}'\mathbf{x} \mid \sum_{j=1}^{n} \tilde{a}_j x_j \le b, \;\; \mathbf{x} \in V \right\}$$

where $\mathbf{c} \in R^{n\times 1}$, and V consists of all constraints whose parameters are certain. $\mathbf{x} \in R^{n\times 1}$ represents a vector of decision variables and $\tilde{\mathbf{a}} \in R^{n\times 1}$ is a vector of uncertain parameters. Let $\mathscr{U}_a$ denote an uncertainty set specified by the modeler. Robust optimization solves an optimization problem assuming that the uncertain data belong to an uncertainty set, $\tilde{\mathbf{a}} \in \mathscr{U}_a$. It looks for an optimal solution that

remains feasible if the uncertainties take any values within that uncertainty set. This reformulation of the problem is referred to as the robust counterpart of the original optimization problem. In some special cases, the robust counterpart of the original problem involves the worst-case outcome of the stochastic data within the uncertainty set, and is a tractable optimization problem with no random parameters.

The robust counterpart of the underlying linear program is formulated as

$$\max_{\mathbf{x}} \min_{\tilde{\mathbf{a}}} \left\{ \mathbf{c}'\mathbf{x} \mid \sum_{j=1}^{n} \tilde{a}_j x_j \leq b,\ \tilde{\mathbf{a}} \in \mathscr{U}_a,\ \mathbf{x} \in V \right\}.$$

The size of the uncertainty set is often related to guarantees on the probability that the constraint involving uncertain coefficients will not be violated. There is a trade-off between the amount of protection against uncertainty that is desired and optimality—the smaller the probability that the constraint will be violated, the more the modeler gives up in terms of optimality of the robust solution relative to the solution to the original optimization problem.

Ellipsoidal, box and polyhedral are the most commonly used uncertainty sets, but more recently, asymmetric uncertainty sets have been used as well in order to capture the probability distribution characteristics of the uncertainties better. In practice, the shape is selected to reflect the modeler's knowledge of the probability distributions of the uncertain parameters, while at the same time making the robust counterpart problem efficiently solvable. Further results on probability bounds related to the size and the shape of uncertainty sets can be found, for example, in Bertsimas and Sim [7] and Bertsimas et al. [8].

For the robust DEA model, we apply an uncertainty set introduced by Bertsimas and Sim [7]. A brief description to this uncertainty set and its adaptation to the DEA modeling is presented next. Let's consider the constraint $\sum_{j=1}^{n} \tilde{a}_j x_j \leq b$ where the uncertain parameter $\tilde{a}_j$ will be robustified. Assume that each entry $\tilde{a}_j$ is modeled by a symmetric and bounded random variable that takes values in $[a_j - \hat{a}_j,\ a_j + \hat{a}_j]$. The random variable $\eta_j = \frac{\tilde{a}_j - \hat{a}_j}{\hat{a}_j}$ which obeys an unknown but symmetric distribution and takes values from an interval $[-1, 1]$. Then the robust counterpart of the linear constraint is derived by a set of the following constraints

$$\left\{ \sum_{j=1}^{n} \hat{a}_j x_j + z\Gamma + \sum_{j=1}^{n} p_j \leq b,\ z + p_j \geq \hat{a}_j t_j, \right.$$
$$\left. -t_j \leq x_j \leq t_j,\ t_j \geq 0,\ p_j \geq 0,\ z \geq 0,\ j = 1, \cdots, n \right\}$$

where the parameter Γ adjusts the robustness of the model against the level of conservatism of the solution. It takes values in the interval $[0, n]$, not necessarily integer. It is crucial to decide the sufficient level where the some parameters are protected to get their worst-case values. When Γ is selected as 0, there is no protection against uncertainty (i.e. uncertainty is ignored). If $\Gamma = n$, then the constraint is completely protected against uncertainty.

14.3.2 Robust DEA Model

Assume that uncertain output parameters, $\tilde{y}_{rj}$ for $r = 1,\cdots,S$, and $j = 1,\cdots,N$, belong to an uncertainty set $\mathscr{U}_y$. The robust counterpart of the DEA model can be formulated as follows;

$$\begin{aligned}
&\max_{\mu,w} \min_{\tilde{y}_{ro}\in\mathscr{U}_y} \sum_{r=1}^{S} \mu_r \tilde{y}_{ro} \\
&\text{s.t.} \quad \sum_{i=1}^{M} w_i x_{io} = 1 \\
&\qquad \min_{\tilde{y}_{rj}\in\mathscr{U}_y} \sum_{r=1}^{S} \mu_r \tilde{y}_{rj} - \sum_{i=1}^{M} w_i x_{ij} \le 0,\ j = 1,\cdots,N,\ \ j \ne o \\
&\qquad \mu_r \ge 0,\ \ w_i \ge 0, \qquad\qquad i = 1,\cdots,M,\ \ r = 1,\cdots,S,
\end{aligned} \tag{14.3}$$

Let θ be a free variable representing the inner minimization problem in the objective function. Then the objective function can be transformed into a constraint

$$\min_{\tilde{y}_{ro}\in\mathscr{U}_y} \sum_{r=1}^{S} \mu_r \tilde{y}_{ro} - \theta \ge 0.$$

Next, to derive the robust counterpart of the DEA model (so-called as the robust DEA model) using the uncertainty set introduced by Bertsimas and Sim [7], both inner minimization problems in the constraints are first solved using dual linear programs. Then the corresponding robust counterparts are reinjected into the corresponding constraints. The robust DEA model can be stated as

$$\begin{aligned}
&\max\ \theta \\
&\text{s.t.} \quad \sum_{i=1}^{M} w_i x_{io} = 1 \\
&\qquad \sum_{r=1}^{S} \mu_r \hat{y}_{ro} - \theta - z_o \Gamma_o - \sum_{r=1}^{S} p_{ro} \ge 0 \\
&\qquad \sum_{i=1}^{M} w_i x_{ij} - \sum_{r=1}^{S} \mu_r \hat{y}_{rj} - \theta - z_j \Gamma_j - \sum_{r=1}^{S} p_{rj} \ge 0,\ j = 1,\cdots,N,\ \ j \ne o \\
&\qquad z_j + p_{rj} \ge \hat{y}_{rj} t_r,\ \ p_{rj} \ge 0,\ \ z_j \ge 0 \qquad\qquad j = 1,\cdots,N,\ r = 1,\cdots,S \\
&\qquad -t_r \le \mu_r \le t_r,\ \ t_r \ge 0, \qquad\qquad r = 1,\cdots,S \\
&\qquad \mu_r \ge 0,\ \ w_i \ge 0, \qquad\qquad i = 1,\cdots,M,\ r = 1,\cdots,S.
\end{aligned} \tag{14.4}$$

where Γ_o and Γ_j represent are the price of robustness of the uncertainty sets defined for the uncertain parameters in the objective function and the constraints, respectively. As explained in detail by Bertsimas and Sim [7] and Sadjadi and Omrani [20], the sufficient level for Γ_j parameter is determined as $\Gamma_j = 1 + \phi^{(-1)}(1 - e_j)\sqrt{N}$ where e_j represents the probability that the constraint j is to be violated, ϕ is the cumulative distribution of standard Gaussian variable. It is also worthwhile to note that the robust DEA model has more variables and constraints than the IDEA model has. On the other hand, it still remains as a tractable linear program.

14.4 Case Study: Performance of Olive Oil Growing Farms

In this study, we consider a real case of olive oil production problem to apply the imprecise and robust DEA models introduced in Sects. 14.2 and 14.3. A sample that consists of 89 olive oil growing farms (labeled as 1–89) in the Aegean region of Turkey is selected to perform an efficiency analysis. The farms are located in the same agricultural area in Izmir; therefore, possible effects of geographical and weather conditions on the oil production are eliminated.

The data were gathered from Taris Olive Oil company, which was established in 2001 by the Union of Taris Olive and Olive Oil Cooperatives. Currently, 33 cooperatives are affiliated with the Union. The company is responsible for the trading of olives cultivated by the olive producing farms located in the Aegean region of Turkey. According to International Olive Council, Turkey produces 4.9 % of world production of olive oil and takes the sixth place in the world olive oil production league (http://www.internationaloliveoil.org/).

Taris has recently started gathering data in order to keep record of the suppliers' performance. The farmers are requested to fill a questionnaire about specifications of farms as well as their performance during the year. For the computational experiments, we use the raw data (relevant to specifications of the farms) that were collected in 2011. Thus, the data set basically reflects the performance of olive oil (not table olives) producers in 2010.

Following to most studies in the literature [3], we also consider total olive land utilized by each farm, cultivation cost, labor as inputs to the DEA models. *Land* is measured by decares ($1\ decare = 1000\,\mathrm{m}^2$). *Cultivation cost* represents the aggregated monetary value in Turkish Liras for cultivation and miscellaneous costs (such as fertilizer, pesticides or fuel costs) spent by the farm in order to operate during a year. We determine *labor* as the number of workers employed to process the harvesting rather than the monetary terms. Apart from land, cost and labor, we also consider *number of olive trees* as the fourth input parameter [2]. Although Land is an input factor, because of the different densities of trees in the given land areas, we also consider number of trees as an input parameter.

As an output parameter, *olive oil yield* is chosen. The olive oil production for each farm depends on the total olive production and is assumed to be uncertain. The other factors such as weather and age of olive trees that might affect the olive production are not taken into account in this study. Since there is no exact measurement for the olive oil production, we consider an expert knowledge in designing the output parameter as the projection of olive production of each farm. Recall that Taris is interested in the olive oil yield rather than the olive production itself.

In current practice, nominal value of olive oil yield for each farm is estimated as 20 % of the olive production. For instance, for a farm that is producing 15,000 kg olives in a year, the nominal olive oil production is expected to be 3000 kg. However, as confirmed by the experts, this value in reality fluctuates within the range of 25 % of the oil production. In this case, the annual olive oil yield varies between 2250 and 3750 for the farm with 3000 kg of nominal olive oil production. Therefore, we develop the imprecise and robust DEA models to evaluate relative

production performance of those olive-growing farms with multiple (deterministic) input parameters (land, cost, labour and number of olive trees) and uncertain output parameter (olive oil production). The deterministic DEA model is used as a benchmark to compare relative performance of those farms using the imprecise and robust DEA models. The DEA models (described in Sects. 14.2 and 14.3) are implemented using General Algebraic Modeling System (GAMS) and solved by a linear programming algorithm.

14.5 Computational Results

We conducted a series of computational experiments to investigate the performance of the imprecise and robust DEA models. Specifically, the experiments aim to answer the following questions:

- How do deterministic, imprecise and robust DEA models perform for the olive oil production problem?
- What are the impact of size and shape of the symmetric uncertainty sets on the robust relative efficiency of olive growing farms?
- How do the robust and imprecise DEA models respond to the changes in size of uncertainty ranges and sample sizes?

In order to measure the relative performance of olive growing farms, we apply the DEA approach. The efficiency scores of farms are computed by solving the linear programs corresponding to the DEA models presented in Sects. 14.2 and 14.3. There is no consensus in the farm efficiency literature on deciding the returns-to-scale assumption. Since all of our farms are located in a small specific region and operate in a similar scale size, constant returns-to-scale is assumed. In summary, we consider the following DEA models;

- *Nominal (deterministic) model* applies the standard DEA approach and uses certain data values of input and output variables. The olive oil production of each farm is calculated as 20 % of the actual olive production.
- *Imprecise DEA approach* considers optimistic and pessimistic views by fixing the upper and lower bounds of the corresponding intervals for the output values in the IDEA model.
- *The robust DEA models* require another input parameter that measures the level of robustness (the price of robustness to use the term from Bertsimas and Sim [7]). The level of robustness varies from 0 to 1.0 and the corresponding DEA models are labelled as $R(0.0), R(0.1), \cdots, R(1.0)$, respectively. In particular, the robust DEA model at zero price of robustness, $R(0.0)$, corresponds to the deterministic (nominal) DEA model.

As mentioned in Sect. 14.3, the robust DEA model requires to specify parameter Γ. In the olive oil production problem, there exists only one uncertain data point at each constraint. The protection level against uncertainty Γ is defined to

vary within interval [0,1]. Since Γ is not necessarily integer, we test the robust DEA model at different protection levels by fixing it at any value within the range [0, 1]. The state $\Gamma = 0$ represents no protection for uncertainty, which corresponds to the nominal model. On the other hand, $\Gamma = 1$ describes a full protection against uncertainty.

14.5.1 Performance Comparison of Imprecise and Robust DEA Approaches

We are first concerned with performance comparisons of all DEA models in terms of relative production efficiency of the farms. Table 14.1 presents the optimal efficiency scores obtained by the deterministic DEA and imprecise DEA approaches in view of optimistic and pessimistic perspectives. In Table 14.2, the results of the robust DEA approach with various level of robustness (at 0.2, 0.4, 0.5, 0.6, 0.8 and 1.0) are summarized in terms of worst-case efficiency scores of farms. The farms taking place on the efficient frontier possess an efficiency score of 1.0 (and highlighted in bold). Notice that the inefficient farms have efficiency scores less than 1.0. Table 14.3 summarizes the statistics of efficiency scores obtained by deterministic, imprecise and robust DEA models in terms of average, minimum and maximum as well as the number of efficient and inefficient farms.

From the results presented in Tables 14.1, 14.2 and 14.3, we can make the following observations;

- The DEA models show different characteristics in terms of the number of efficient farms. The optimistic (pessimistic) strategy obtained by IDEA provides 32 farms efficient while the deterministic DEA produces only 8 efficient farms. On the other hand, no farm is declared as efficient according to the robust DEA strategy. Only farm 42 (see Table 14.2) is determined as perfectly efficient since it is declared as efficient by all nominal, optimistic and pessimistic imprecise DEA models.
- The IDEA model with optimistic approach seems the least conservative way of evaluation in the presence of uncertain data in the form of bounds. The efficiency scores obtained by the optimistic model are consistently greater than or equal to the efficiency scores produced by the nominal DEA and pessimistic IDEA models. As mentioned in Sect. 14.2, the IDEA in view of an optimistic perspective assumes that a farm under consideration performs at its best production whereas the rest of farms perform at their worst production efficiency. The highest average efficiency score is achieved by the optimistic IDEA as 74.4 %. As expected, it outperforms the deterministic DEA strategy that, in average, provides 52.4 % of overall scores.
- On the other hand, the IDEA model with pessimistic perspective is seen the most conservative way of finding efficiency scores of olive growing farms as it produces the lowest efficiency scores comparing with the nominal DEA and the

Table 14.1: Efficiency scores obtained by the deterministic and imprecise DEA models

	Deterministic	IDEA			Deterministic	IDEA	
Farms	DEA	Optimistic	Pessimistic	Farms	DEA	Optimistic	Pessimistic
1	**1.00**	**1.00**	0.63	46	0.25	0.41	0.15
2	0.22	0.36	0.13	47	0.27	0.44	0.16
3	0.42	0.70	0.25	48	0.27	0.45	0.16
4	0.07	0.11	0.04	49	0.16	0.27	0.10
5	**1.00**	**1.00**	0.77	50	0.26	0.43	0.16
6	0.34	0.57	0.20	51	0.21	0.35	0.13
7	0.45	0.75	0.27	52	0.48	0.80	0.29
8	0.20	0.34	0.12	53	0.38	0.64	0.23
9	0.24	0.41	0.15	54	0.35	0.58	0.21
10	0.10	0.16	0.06	55	0.59	0.99	0.36
11	0.50	0.83	0.30	56	0.21	0.36	0.13
12	0.56	0.93	0.33	57	0.79	**1.00**	0.48
13	0.52	0.87	0.31	58	0.37	0.62	0.22
14	0.34	0.57	0.21	59	0.42	0.70	0.25
15	0.12	0.20	0.07	60	0.66	**1.00**	0.39
16	**1.00**	**1.00**	0.91	61	0.37	0.62	0.22
17	0.72	**1.00**	0.43	62	0.38	0.64	0.23
18	0.67	**1.00**	0.40	63	0.47	0.79	0.28
19	**1.00**	**1.00**	0.99	64	0.90	**1.00**	0.54
20	0.73	**1.00**	0.44	65	0.76	**1.00**	0.46
21	0.65	**1.00**	0.39	66	0.50	0.83	0.30
22	0.67	**1.00**	0.40	67	0.71	**1.00**	0.43
23	0.50	0.83	0.30	68	0.58	0.97	0.35
24	**1.00**	**1.00**	0.75	69	0.81	**1.00**	0.49
25	0.33	0.54	0.20	70	**1.00**	**1.00**	0.85
26	0.56	0.94	0.34	71	0.76	**1.00**	0.46
27	0.97	**1.00**	0.58	72	0.51	0.86	0.31
28	0.10	0.16	0.06	73	0.59	0.98	0.35
29	0.58	0.96	0.35	74	0.75	**1.00**	0.45
30	0.74	**1.00**	0.45	75	0.55	0.91	0.33
31	0.19	0.31	0.11	76	0.31	0.52	0.19
32	0.34	0.56	0.20	77	0.37	0.62	0.22
33	0.99	**1.00**	0.59	78	0.95	**1.00**	0.57
34	**1.00**	**1.00**	0.82	79	0.44	0.74	0.27
35	0.19	0.32	0.11	80	0.31	0.51	0.18
36	0.19	0.31	0.11	81	0.66	**1.00**	0.40
37	0.71	**1.00**	0.43	82	0.82	**1.00**	0.49
38	0.08	0.13	0.05	83	0.50	0.83	0.30
39	0.42	0.69	0.25	84	0.30	0.49	0.18
40	0.56	0.93	0.34	85	0.58	0.97	0.35
41	0.51	0.86	0.31	86	0.47	0.79	0.28
42	**1.00**	**1.00**	**1.00**	87	0.80	**1.00**	0.48
43	0.14	0.23	0.08	88	0.60	**1.00**	0.36
44	0.64	**1.00**	0.38	89	0.69	**1.00**	0.42
45	0.32	0.53	0.19				

Table 14.2: Efficiency scores obtained by the robust DEA model at various level of robustness

Farms	Robust DEA R(0.0)	R(0.2)	R(0.4)	R(0.5)	R(0.6)	R(0.8)	R(1.0)	Farms	Robust DEA R(0.0)	R(0.2)	R(0.4)	R(0.5)	R(0.6)	R(0.8)	R(1.0)
1	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60	46	0.25	0.22	0.20	0.19	0.18	0.17	0.15
2	0.22	0.20	0.18	0.17	0.16	0.14	0.13	47	0.27	0.24	0.22	0.21	0.20	0.18	0.16
3	0.42	0.38	0.34	0.33	0.31	0.28	0.25	48	0.27	0.24	0.22	0.21	0.20	0.18	0.16
4	0.07	0.06	0.06	0.05	0.05	0.05	0.04	49	0.16	0.15	0.13	0.13	0.12	0.11	0.10
5	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60	50	0.26	0.23	0.21	0.20	0.19	0.17	0.16
6	0.34	0.31	0.28	0.27	0.25	0.23	0.20	51	0.21	0.19	0.17	0.16	0.16	0.14	0.13
7	0.45	0.41	0.37	0.35	0.33	0.30	0.27	52	0.48	0.44	0.40	0.38	0.36	0.32	0.29
8	0.20	0.19	0.17	0.16	0.15	0.14	0.12	53	0.38	0.35	0.31	0.30	0.28	0.26	0.23
9	0.24	0.22	0.20	0.19	0.18	0.16	0.15	54	0.35	0.32	0.29	0.27	0.26	0.23	0.21
10	0.10	0.09	0.08	0.07	0.07	0.06	0.06	55	0.59	0.54	0.48	0.46	0.44	0.39	0.36
11	0.50	0.45	0.41	0.39	0.37	0.33	0.30	56	0.21	0.19	0.18	0.17	0.16	0.14	0.13
12	0.56	0.50	0.46	0.43	0.41	0.37	0.33	57	0.79	0.72	0.65	0.62	0.59	0.53	0.48
13	0.52	0.47	0.43	0.41	0.39	0.35	0.31	58	0.37	0.34	0.31	0.29	0.28	0.25	0.22
14	0.34	0.31	0.28	0.27	0.25	0.23	0.21	59	0.42	0.38	0.34	0.33	0.31	0.28	0.25
15	0.12	0.11	0.10	0.09	0.09	0.08	0.07	60	0.66	0.59	0.54	0.51	0.49	0.44	0.39
16	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60	61	0.37	0.34	0.31	0.29	0.28	0.25	0.22
17	0.72	0.65	0.59	0.56	0.53	0.48	0.43	62	0.38	0.35	0.31	0.30	0.28	0.26	0.23
18	0.67	0.61	0.55	0.52	0.50	0.45	0.40	63	0.47	0.43	0.39	0.37	0.35	0.31	0.28
19	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60	64	0.90	0.82	0.74	0.70	0.67	0.60	0.54
20	0.73	0.66	0.60	0.57	0.54	0.49	0.44	65	0.76	0.69	0.63	0.59	0.56	0.51	0.46
21	0.65	0.59	0.53	0.51	0.48	0.43	0.39	66	0.50	0.45	0.41	0.39	0.37	0.33	0.30
22	0.67	0.61	0.55	0.52	0.50	0.45	0.40	67	0.71	0.64	0.58	0.55	0.53	0.47	0.43
23	0.50	0.45	0.41	0.39	0.37	0.33	0.30	68	0.58	0.52	0.47	0.45	0.43	0.39	0.35
24	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60	69	0.81	0.73	0.66	0.63	0.60	0.54	0.49
25	0.33	0.29	0.27	0.25	0.24	0.22	0.20	70	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60
26	0.56	0.51	0.46	0.44	0.42	0.38	0.34	71	0.76	0.69	0.62	0.59	0.56	0.51	0.46
27	0.97	0.87	0.79	0.75	0.71	0.64	0.58	72	0.51	0.47	0.42	0.40	0.38	0.34	0.31
28	0.10	0.09	0.08	0.07	0.07	0.06	0.06	73	0.59	0.53	0.48	0.46	0.44	0.39	0.35
29	0.58	0.52	0.47	0.45	0.43	0.39	0.35	74	0.75	0.68	0.61	0.58	0.55	0.50	0.45

Table 14.2: (continued)

Farms	Robust DEA R(0.0)	R(0.2)	R(0.4)	R(0.5)	R(0.6)	R(0.8)	R(1.0)	Farms	Robust DEA R(0.0)	R(0.2)	R(0.4)	R(0.5)	R(0.6)	R(0.8)	R(1.0)
30	0.74	0.67	0.61	0.58	0.55	0.50	0.45	75	0.55	0.49	0.45	0.42	0.40	0.36	0.33
31	0.19	0.17	0.15	0.14	0.14	0.12	0.11	76	0.31	0.28	0.26	0.25	0.23	0.21	0.19
32	0.34	0.31	0.28	0.26	0.25	0.23	0.20	77	0.37	0.34	0.30	0.29	0.27	0.25	0.22
33	0.99	0.90	0.81	0.77	0.73	0.66	0.59	78	0.95	0.86	0.78	0.74	0.70	0.63	0.57
34	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60	79	0.44	0.40	0.36	0.34	0.33	0.30	0.27
35	0.19	0.17	0.15	0.15	0.14	0.13	0.11	80	0.31	0.28	0.25	0.24	0.23	0.20	0.18
36	0.19	0.17	0.15	0.14	0.14	0.12	0.11	81	0.66	0.60	0.54	0.51	0.49	0.44	0.40
37	0.71	0.65	0.59	0.56	0.53	0.48	0.43	82	0.82	0.74	0.67	0.64	0.61	0.55	0.49
38	0.08	0.07	0.07	0.06	0.06	0.05	0.05	83	0.50	0.45	0.41	0.39	0.37	0.33	0.30
39	0.42	0.38	0.34	0.32	0.31	0.28	0.25	84	0.30	0.27	0.24	0.23	0.22	0.20	0.18
40	0.56	0.51	0.46	0.44	0.41	0.37	0.34	85	0.58	0.53	0.48	0.45	0.43	0.39	0.35
41	0.51	0.46	0.42	0.40	0.38	0.34	0.31	86	0.47	0.43	0.39	0.37	0.35	0.32	0.28
42	**1.00**	0.91	0.82	0.78	0.74	0.67	0.60	87	0.80	0.72	0.65	0.62	0.59	0.53	0.48
43	0.14	0.12	0.11	0.11	0.10	0.09	0.08	88	0.60	0.55	0.49	0.47	0.45	0.40	0.36
44	0.64	0.58	0.52	0.50	0.47	0.42	0.38	89	0.69	0.63	0.57	0.54	0.51	0.46	0.42
45	0.32	0.29	0.26	0.25	0.23	0.21	0.19								

optimistic IDEA models. This case also confirms the findings in [19]. The average efficiency score is 33.6 % for the pessimistic model and 88 farms are found to be inefficient. The discrimination of efficiency scores exhibits a noticeable change between optimistic to pessimistic modeling.

- Under the full protection against uncertainty, the average efficiency score for the farms is obtained as 31.5 %. This is slightly lower than the average score for the pessimistic IDEA model, which yields 33.6 % average efficiency. Note that 81 out of 89 farms have protected their efficiency scores at the same level with the pessimistic IDEA model. On the other hand, in case of no protection against uncertainty, not surprisingly, the model produces the same efficiency scores with the nominal (deterministic) model (an average of 52.4 %). This verifies the efficiency scores obtained by the deterministic model.

Table 14.3: Statistics of efficiency scores

	Deterministic	Imprecise DEA		Robust DEA					
	DEA	Optimistic	Pessimistic	R(0.2)	R(0.4)	R(0.5)	R(0.6)	R(0.8)	R(1.0)
Efficiency scores (%)									
Average	52.4	74.4	33.6	47.4	42.9	40.8	38.8	35.0	31.5
Min	6.8	11.3	4.1	6.1	5.5	5.3	5.0	4.5	4.1
Max	100	100	100	90.5	81.8	77.8	73.9	66.7	60.0
Number of efficient and inefficient farms									
Efficient	8	32	1	0	0	0	0	0	0
Inefficient	81	57	88	89	89	89	89	89	89

- No farm is reported as efficient by the robust DEA approach. For instance, the robust DEA model with 50 % protection against the uncertainty produces the maximum efficiency score as 77.8 % and the average efficiency score is 40.8 %. Using this reference efficiency score obtained by robust DEA model (at 0.5 price of robustness for all constraints), we can compare the relative production performance of the DEA models. We observe that

  – The efficiency scores of all farms obtained by the robust DEA ($E_{R(0.50)}$) are always lower than those scores achieved by the nominal model (E_N) and the robust DEA model with no protection ($E_{R(0.0)}$).
  – Their efficiency scores ($E_{R(0.50)}$) are persistently larger than those achieved by the pessimistic imprecise DEA (E_{pes}) and the robust DEA model with full protection against uncertainty ($E_{R(1.0)}$).
  – On the other hand, the imprecise DEA scores under the optimistic view (E_{opt}) are grater than the scores of all other DEA approaches.

  As a result, we can state the following relationship between the efficiency scores of all farms obtained by various DEA approaches as

$$E_{R(1.0)} \leq E_{pes} \leq E_{R(0.50)} \leq E_N = E_{R(0.0)} \leq E_{opt}.$$

- Finally, we can comment on the impact of level of conservatism on the average efficiency scores obtained by the robust DEA models at various price of robustness. As the level of conservatism increases from no-protection to full-protection against uncertainty (i.e. varying price of robustness within [0, 1]), the average efficiency scores for the sample decline from 52.4 to 31.5 %. Similarly, the lowest (highest) average efficiency scores decrease from 6.8 (100) to 4.1 % (60 %). This basically shows a trade-off between the level conservatism and the production efficiency level that a decision maker needs to take into account.

14.5.2 Impact of Model Parameters

We also investigate possible impact of the model parameters and the size of future uncertainty on imprecise and robust production performance of olive oil producing farms. We adopt the simulation framework suggested by Shokouhi et al. [21] for the case study. The olive-growing farms are clustered into smaller groups according to their efficiency scores obtained by the nominal DEA approach. In this chapter, due to the length restriction, we only present the results of three groups each of whom consists of eight farms with the same or similar nominal performance. The farms in *Group 1* are all declared as efficient whereas *Group 2* consists of the least efficient farms according to the nominal DEA approach. *Group 3* involves such farms showing medium level (around 50 %) of efficiency.

We design two experiments, labeled as *Experiment I* and *Experiment II*, with different size of uncertainty levels for the output parameters. More precisely, *Experiment I* assumes the initial range for the uncertain olive oil yield of a farm where the olive oil production of a farm deviates from its nominal production by 1/4 (25 %). On the other hand, *Experiment II* is designed to observe the impact of the interval size on results by considering narrower ranges for olive oil yields of the farms where the olive oil production of a farm is assumed to deviate from its nominal production by 1/6 (approximately 16 %) rather than 1/4.

A brief description of the simulation procedure is as follows. First, we calculate the optimal weights associated with each farm using the nominal values of input and the estimated output parameters within the imprecise and robust DEA models (given the price of robustness). Secondly, we randomly generated 1000 nominal values of olive oil production levels (using uniform distribution) for the uncertain output parameters. Thus, the corresponding intervals for each simulated point within the ranges of 25 and 16 % olive production are then determined. Finally, the optimal weights (obtained with the estimated oil production level) and oil olive production intervals at each generated random points of the output parameters are then used to find the robust efficiency scores by solving the robust DEA models. The same procedure is repeated with 1000 output intervals in the same manner. We then analyse the statistics of 1000 efficiency scores associated with each farm within three groups in terms of average and standard deviation of scores.

Table 14.4 illustrates the simulation results for Experiment I (top) and Experiment II (bottom) using three groups of farms in terms of average efficiency scores obtained at various price of robustness with robust DEA and IDEA with optimistic (IDEA-O) and pessimistic (IDEA-P) approaches. These results basically show, on average, how the olive oil production performance of each farm changes under imprecise and robust DEA when the range of uncertainty varies.

From the simulation results in Table 14.4, we observe that

- The average efficiency scores in Experiment II are always higher than those in Experiment I regardless the choice of models (imprecise and robust DEA at each price of robustness). This implies that as the size of intervals for the random parameters decreases (i.e. random parameter approaches to the estimated nominal value), the average efficiency score increases.
- As the simulation results confirm, the imprecise DEA with optimistic view produces the highest average efficiency scores (labelled as SE), obtained out of 1000 simulated points for all farms. In both Experiments I and II, as the price of robustness varies between 0 and 1, the average efficiency score decreases. As a result, we can state the following relationship between the efficiency scores of all farms obtained by various DEA approaches as

$$SE_{R(1.0)} \leq \cdots \leq SE_{R(0.0)} \leq SE_{opt}.$$

- On the other hand, the IDEA approach with pessimistic view produces the lowest average efficiency scores (obtained by the simulation experiments) for most farms in three groups. The lowest scores are indicated in bold in Table 14.4. Therefore, $SE_{pes} \leq SE_{R(1.0)}$. Notice that the order (between the IDEA approach with pessimistic view and the most conservative robust optimization approach) that was already established from Tables 14.1 and 14.2 has changed. For these cases, IDEA apparently becomes more conservative than robust DEA. This result leads us to conclude that for such homogeneous smaller samples of farms (in groups 1, 2 and 3) with similar nominal performance, the pessimistic IDEA approach produces more conservative scores than the robust DEA model with full protection no matter which uncertainty range (25 or 16 %) is chosen. However, for more diversified sample of farms, the robust DEA strategy is more conservative than IDEA with pessimistic view as illustrated in Tables 14.2 and 14.3.
- In order to illustrate the overall performance (labeled as "AverGP" in Table 14.4) of each farm within the three different groups, we compute the average of the efficiency scores obtained by the robust DEA at eleven values of price of robustness $(0, 0.1, 0.2, \cdots, 1.0)$. One can easily see that all farms in both Groups 1 and 3 exhibit average performance around 80 and 85 % in Experiments I and II, respectively. However, farms in Group 2 show different average performance varying from 42 to 81 % in Experiment I (and similarly, 45–85 % in Experiment II). Recall that these farms are reported as the least efficient by the deterministic DEA model.

Table 14.4: Expected efficiency scores (%) obtained by the simulation experiments

	Group 1								Group 2								Group 3							
Farm ID	1	5	16	19	24	34	42	70	4	10	15	28	31	38	43	49	11	13	23	41	52	66	72	83
Approach	*Experiment I: Range of 25 % of the oil production*																							
IDEA-O	90	87	91	89	91	88	88	88	80	90	90	88	94	90	90	91	94	86	87	88	93	88	85	91
IDEA-P	**57**	78	79	88	**68**	73	88	80	**29**	**57**	**47**	**34**	88	**35**	**52**	77	**64**	74	78	**56**	**53**	70	**58**	**55**
R(0.0)	88	89	88	91	89	88	90	88	48	90	78	56	94	58	87	90	86	87	89	86	86	86	90	87
R(0.2)	84	85	85	86	86	84	86	84	46	86	74	53	86	55	81	88	84	84	85	83	83	82	86	83
R(0.4)	81	81	81	81	81	81	81	81	43	82	70	50	81	53	77	83	81	82	80	80	80	82	79	80
R(0.5)	79	78	79	79	79	79	80	78	43	79	68	49	79	51	76	79	80	78	78	78	78	78	80	78
R(0.6)	77	77	78	77	77	77	78	77	41	77	66	48	77	50	74	78	77	77	77	76	76	77	78	77
R(0.8)	73	73	74	74	73	73	73	74	39	73	63	46	73	48	69	74	73	73	74	74	73	73	73	74
R(1.0)	70	**70**	**70**	**70**	70	**69**	**70**	**71**	37	70	60	43	**70**	45	66	**71**	70	**70**	**70**	70	70	**69**	70	71
AverGP	79	79	80	80	79	79	80	79	42	80	68	49	80	51	76	81	79	79	79	78	78	79	79	78
Approach	*Experiment II: Range of 16 % of the oil production*																							
IDEA-O	97	95	96	96	96	96	96	95	77	96	96	88	96	91	96	95	98	95	95	95	95	96	95	95
IDEA-P	**67**	**67**	**67**	**73**	**67**	**68**	**67**	**67**	**35**	**67**	**56**	**40**	89	**41**	**62**	88	**75**	87	86	**68**	**64**	84	**69**	**67**
R(0.0)	92	93	92	94	92	92	93	93	50	93	80	57	96	60	88	93	91	91	93	90	91	93	93	92
R(0.2)	89	90	89	90	90	89	90	89	48	90	77	56	90	58	85	92	89	90	89	88	88	88	90	88
R(0.4)	86	87	87	87	86	87	87	87	46	87	74	53	87	57	82	89	86	87	86	86	86	88	87	86
R(0.5)	85	85	85	86	85	85	86	85	46	86	73	52	85	55	81	85	86	85	85	85	84	84	84	85
R(0.6)	84	84	85	84	84	84	84	84	45	84	72	52	84	55	80	84	84	83	84	83	83	84	83	84
R(0.8)	81	81	82	81	81	81	81	81	41	78	67	48	78	50	73	78	81	81	81	81	81	81	81	81
R(1.0)	78	79	79	78	79	78	78	79	39	73	62	45	**73**	47	69	**74**	78	**79**	**79**	78	79	**78**	78	79
AverGP	85	85	86	86	85	85	86	85	45	84	72	52	84	54	80	85	85	85	85	85	84	85	85	85

14.6 Conclusions

In this study, we are concerned with performance evaluation of olive oil production using DEA under uncertainty. In particular, we study the sensitivity of efficiency scores obtained through imprecise and robust optimization based DEA approaches in a real world agricultural problem. The olive oil production problem involves the efficiency assessment of a sample of farms located in a specific region in Turkey. The only output factor (olive oil yield of the farm) is uncertain varying between bounds that depend on the olive production. For computational experiments, we implement two basic approaches of imprecise DEA and robust optimization based DEA models.

The results indicate that the optimistic model yields higher levels of efficiency for the farms, whereas the pessimistic model scores are way below than the optimistic and nominal models as expected. The discrimination of the scores is considerably worse in the pessimistic model where only one farm remains efficient. In robust DEA modeling, as the level of conservatism increases from no-protection to full-protection against uncertainty, the average, minimum and maximum efficiency scores for the sample decline. This indicates a trade-off between the level conservatism and the efficiency levels.

We compare the efficiency scores of the DEA models in order to establish performance ranking of deterministic, imprecise and robust DEA approaches. The IDEA with optimistic view yields considerably higher levels of efficiency than any other DEA models considered in this study. When no robustness is assumed, the efficiency scores are exactly the same with those of the nominal model. Under full robustness, the efficiency scores of robust DEA are less than or equal to the scores of the IDEA with pessimistic view. Therefore, it can be stated that the most conservative robust DEA model can yield lower efficiency scores than the most pessimistic imprecise DEA model.

In order to measure sensitivity of different DEA approaches to changing uncertainty ranges and parameters, we perform simulation based experiments using homogeneous groups of farms. The simulation results reveal that when the uncertainty ranges are close to the estimated nominal values, the average efficiency scores increase. In addition, when the level of conservatism increases from no-protection to full-protection against uncertainty, the average efficiency scores for the sample decline. Therefore, we can conclude that the choice of price of robustness and size of intervals play an important role on the performance of the robust DEA models. As future research directions, one may investigate model behavior in cases where uncertainty is observed in both input and output parameters simultaneously. In particular, data driven uncertainty sets would be worthwhile to investigate.

Acknowledgements The authors are grateful to Isil Cillidag and Asli Cetiner from Taris Zeytin Company, Izmir, Turkey for their collaboration and support in data collection for this research.

References

1. Alcaide-López-de-Pablo, D., Dios-Palomares, R., Prieto, Á.M.: A new multicriteria approach for the analysis of efficiency in the Spanish olive oil sector by modelling decision maker preferences. Eur. J. Oper. Res. **234**(1), 241–252 (2014)
2. Amores, A.F., Contreras, I.: New approach for the assignment of new European agricultural subsidies using scores from data envelopment analysis: application to olive-growing farms in andalusia (Spain). Eur. J. Oper. Res. **193**(3), 718–729 (2009)
3. Atici, K.B.: Using data envelopment analysis for the efficiency and elasticity evaluation of agricultural farms. Ph.D. thesis, University of Warwick (2012)
4. Ben-Tal, A., Nemirovski, A.: Robust solutions of uncertain linear programs. Oper. Res. Lett. **25**(1), 1–13 (1999)
5. Ben-Tal, A., Nemirovski, A.: Robust solutions of linear programming problems contaminated with uncertain data. Math. Program. **88**(3), 411–424 (2000)
6. Ben-Tal, A., Ghaoui, L.E., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
7. Bertsimas, D., Sim, M.: The price of robustness. Oper. Res. **52**(1), 35–53 (2004)
8. Bertsimas, D., Pachamanova, D., Sim, M.: Robust linear optimization under general norms. Oper. Res. Lett. **32**, 510–516 (2004)
9. Blancard, S., Boussemart, J.P., Leleu, H.: Measuring potential gains from specialization under non-convex technologies. J. Oper. Res. Soc. **62**(10), 1871–1880 (2011)
10. Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. Eur. J. Oper. Res. **2**(6), 429–444 (1978)
11. Cooper, W.W., Park, K.S., Yu, G.: IDEA and AR-IDEA: models for dealing with imprecise data in DEA. Manag. Sci. **45**(4), 597–607 (1999)
12. Cooper, W.W., Park, K.S., Yu, G.: An illustrative application of IDEA (imprecise data envelopment analysis) to a Korean mobile telecommunication company. Oper. Res. **49**(6), 807–820 (2001)
13. Cooper, W.W., Seiford, L.M., Tone, K.: Introduction to Data Envelopment Analysis and Its Uses. Springer, New York (2006)
14. Despotis, D.K., Smirlis, Y.G.: Data envelopment analysis with imprecise data. Eur. J. Oper. Res. **140**(1), 24–36 (2002)
15. Dios-Palomares, R., Martínez-Paz, J.M.: Technical, quality and environmental efficiency of the olive oil industry. Food Policy **36**(4), 526–534 (2011)
16. Ghaoui, L.E., Lebret, H.: Robust solutions to least-squares problems with uncertain data. SIAM J. Matrix Anal. Appl. **18(4)**, 1035–1064 (1997)
17. Kim, S.H., Park, C.G., Park, K.S.: An application of data envelopment analysis in telephone offices evaluation with partial data. Comput. Oper. Res. **26**(1), 59–72 (1999)
18. Liu, J.S., Lu, L.Y., Lu, W.M., Lin, B.J.: A survey of DEA applications. Omega **41**(5), 893–902 (2013)
19. Park, K.: Efficiency bounds and efficiency classifications in DEA with imprecise data. J. Oper. Res. Soc. **58**(4), 533–540 (2007)
20. Sadjadi, S., Omrani, H.: Data envelopment analysis with uncertain data: an application for Iranian electricity distribution companies. Energy Policy **36**(11), 4247–4254 (2008)
21. Shokouhi, A.H., Hatami-Marbini, A., Tavana, M., Saati, S.: A robust optimization approach for imprecise data envelopment analysis. Comput. Ind. Eng. **59**(3), 387–397 (2010)
22. Soyster, A.L.: Convex programming with set-inclusive constraints and application to inexact linear programming. Oper. Res. **21**(5), 1154–1157 (1973)
23. Thanassoulis, E.: Introduction to the Theory and Application of Data Envelopment Analysis. Kluwer Academic Publishers, Dordrecht (2001)
24. Zhu, J.: Imprecise data envelopment analysis (IDEA): a review and improvement with an application. Eur. J. Oper. Res. **144**(3), 513–529 (2003)
25. Zhu, J.: Imprecise DEA via standard linear DEA models with a revisit to a Korean mobile telecommunication company. Oper. Res. **52**(2), 323–329 (2004)

Index

M. Doumpos et al. (eds.), *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science 241, DOI 10.1007/978-3-319-33121-8