Computers & Operations Research 66 (2016) 181–189

Contents lists available at ScienceDirect

Computers & Operations Research

journal homepage: www.elsevier.com/locate/caor

A bicriteria approach to robust optimization ☆

André Chassein *, Marc Goerigk

Fachbereich Mathematik, Technische Universität Kaiserslautern, Germany

ARTICLE INFO

Available online 5 September 2015

Keywords:
Robust optimization
Column generation
Minimum cost flow problem
Bicriteria optimization
Linear programming

ABSTRACT

The classic approach in robust optimization is to optimize the solution with respect to the worst case scenario. This pessimistic approach yields solutions that perform best if the worst scenario happens, but also usually perform bad for an average case scenario. On the other hand, a solution that optimizes the performance of this average case scenario may lack in the worst-case performance guarantee.

In practice it is important to find a good compromise between these two solutions. We propose to deal with this problem by considering it from a bicriteria perspective. The Pareto curve of the bicriteria problem visualizes exactly how costly it is to ensure robustness and helps to choose the solution with the best balance between expected and guaranteed performance.

In this paper we consider linear programming problems with uncertain cost functions. Building upon a theoretical observation on the structure of Pareto solutions for these problems, we present a column generation approach that requires no direct solution of the computationally expensive worst-case problem. In computational experiments we demonstrate the effectiveness of both the proposed algorithm, and the bicriteria perspective in general.

1. Introduction

Robust and stochastic optimization are paradigms for optimization under uncertainty, that have been receiving increasing attention over the last two decades (see the recent textbook [3]). Optimization under uncertainty means that the exact parameters that describe the optimization problem are not known exactly and can only be estimated. In contrast to stochastic optimization, where one assumes to have enough knowledge to estimate the probability distribution of the input data, robust optimization deals with problems without any or only very little information about the underlying distributions. In robust optimization one defines an uncertainty set that describes all possible realizations of the input data. This can be done, for example, by defining a finite set of different scenarios for the parameter values, but also continuous uncertainty sets are possible. The aim is to find a solution that is feasible for all realization of input data and also yields the best performance if the worst possible realization occurs.

This approach follows a pessimistic point of view and, hence, it is not surprising that optimizing only the worst case performance yields in most cases a solution that performs poorly in the average case, which makes the solution impractical for many applications. Even though it may not be clear what the average realization of the data is, as there is no information about the distribution of data available, a lot of effort has been put into the development of robustness concepts that reduce the conservatism of the solution and give a better performance in the average case.

Several such approaches to overcome this conservatism have been proposed. Following the ideas of Ben-Tal and Nemirovski [4] it is a matter of choosing the right uncertainty set to get a solution that performs well in the average case and in the worst case. Bertsimas and Sim [8] introduce a parameter Γ that allows controlling the conservatism of a solution. Fischetti and Monaci [11] propose to identify a nominal scenario and to demand from the robust solution a performance guarantee for this scenario. For general surveys on robust optimization, we refer to [12,2,3,7]. In this paper, we take a simpler and more direct approach to relax the conservatism of a solution: We propose to include the average case performance as an objective function, thus resulting in a bicriteria optimization problem.

A frequently used assumption is that the objective function of the optimization problem is certain, as every objective function can be represented as a constraint by using the epigraph transformation. While this is a valid method, it has the drawback that feasibility and performance guarantee are mixed into one criterion, which is questionable for most practical problems. Therefore,

☆ 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.

* Corresponding author.
E-mail address: chassein@mathematik.uni-kl.de (A. Chassein).

http://dx.doi.org/10.1016/j.cor.2015.08.007
0305-0548/

we focus in this paper explicitly on problems that are affected by uncertainty only in the objective function. This can be done by restricting the set of feasible solutions to solutions that are feasible for all possible parameter values. We define the bicriteria optimization problem with the two objective functions average and worst case performance, and the Pareto front of this problem as the *average case–worst case curve* (*AC–WC curve*). We argue that the AC–WC curve is a valuable tool to assess the trade-off between average and worst-case performance, and should play a vital role within a robust decision making process. Note that we do not start with an uncertain multicriteria optimization problem, as considered in [13,10]. Instead, we begin with an uncertain single criterion optimization problem and extend it to a robust bicriteria optimization problem.

To compute the AC–WC curve, we make algorithmic use of the observation that the average and worst case performance of a robust solution can be interpreted as a special point on the Pareto front of a multicriteria optimization problem where every possible scenario outcome leads to its own objective (see [15,2]). To the best of our knowledge, this is the first time that this observation is used.

The paper is organized as follows. In Section 2 we introduce basic definitions and notations that are used throughout the paper. In Section 3 we show a theoretical result that allows us to develop a column generation approach to compute the AC–WC curve. We evaluate different experiments in Section 3.1. The first experiment compares the newly developed column generation approach to compute the AC–WC curve with a straightforward approach. The second experiment uses an approximation algorithm from the literature to approximate the AC–WC curve. The gains that can be obtained by using the AC–WC curve instead of only considering the average and worst case solutions are shown in the third experiment, and in the last experiment we use the AC–WC curve to directly compare the performance of two frequently used robustness concepts from the literature. For all experiments we use the minimum cost flow problem as a benchmark. We conclude the paper and point to further research questions in Section 3.2.

2. Notation and definitions

We consider single criterion linear programming problems (LP) of the form:

$$\begin{aligned} \min\quad & c^T x \\ \text{s. t.}\quad & Ax = b \\ & x \ge 0, \end{aligned} \tag{LP}$$

where $x \in \mathbb{R}^N$ are the decision variables. Such a problem can be solved by general linear programming solvers; however, for many problems, (e.g., the minimum cost flow problem, the maximum flow problem, or the transportation problem), there exist specialized algorithms that outperform general linear programming solvers. Note that these specialized algorithms usually need the exact structure of problem (LP).

The uncertainty is introduced by an uncertainty set $\mathcal{U}$. A frequently used uncertainty set is the interval uncertainty, which is given as a hyper-rectangle $\mathcal{U} = \times_{i=1}^N [\underline{c}_i, \overline{c}_i]$. It is a standard assumption that the average scenario is given by the midpoint of the interval $\hat{c} = 0.5(\underline{c} + \overline{c})$. A second important kind of uncertainty are discrete uncertainty sets of the form $\mathcal{U} = \{c_1, \dots, c_n\}$ that specify n different cost vectors for the objective function for each scenario. Denote by $\hat{c} = \frac{1}{n}\sum_{i=1}^n c_i$ the average cost vector, assuming a uniform probability distribution. If any other probability distribution p is available, where p_i is the probability that scenario i is realized, and one is interested in optimizing the expected value, the cost vector $\hat{c}(p) = \sum_{i=1}^n p_i c_i$ can be used instead. As there is no structural difference between these two cases, we will deal in the following only with vectors $\hat{c}$ that implicitly assume that every scenario is equally likely. We assume in this section to have discrete uncertainty to be able to exploit the discrete structure of $\mathcal{U}$. Using this notation, the average case optimization problem (AC) has the form:

$$\begin{aligned} \min\quad & \hat{c}^T x \\ \text{s. t.}\quad & Ax = b \\ & x \ge 0. \end{aligned} \tag{AC}$$

Note that solving problem (AC) has the same computational complexity as solving the original problem (LP). In general, this does not hold for the worst-case (robust) optimization problem (WC) that looks as follows:

$$\begin{aligned} \min\quad & z \\ \text{s. t.}\quad & c_k^T x \le z \quad k = 1, \dots, n \\ & Ax = b \\ & x \ge 0 \end{aligned} \tag{WC}$$

Both problems (AC) and (WC) are tractable, as they are formulated as linear programs. Both optimization goals – the average case as well as the worst case – are important criteria to evaluate. But in most cases these two functions are contradicting. A good performance in the average case often has to be paid with worse performance in a single scenario and vice versa: good performance in the worst case objective leads to a bad performance in the average case. A common approach to deal with contradicting objective functions is to translate the problem into a multicriteria optimization problem. Applied to our situation this yields a bicriteria optimization problem $(BILP)$ with the two objective functions average and worst case performance

$$\begin{aligned} vec-\min\quad & (z, \hat{c}^T x) \\ \text{s. t.}\quad & c_k^T x \le z \qquad k = 1, \dots, n \\ & Ax = b \\ & x \ge 0 \end{aligned} \tag{BILP}$$

Note that in most cases there is not a single solution that optimizes problem $(BILP)$; on the contrary, there may be many solutions that can be seen as optimal solutions for problem $(BILP)$. A common approach to solve such a multicriteria optimization problem is to compute the set of all solutions that are Pareto efficient. A solution x is called Pareto efficient if there exists no solution y that performs as least as well as x in every objective function and is strictly better in at least one objective function. The Pareto front of a multicriteria optimization problem is obtained by mapping all Pareto efficient solutions of the problem into the objective space with the corresponding objective functions. We define the *AC–WC curve* as the Pareto front of problem $(BILP)$.

Example 2.1. We illustrate the concept of the AC–WC curve using the following uncertain linear program that can be seen as a diversification problem

$$\begin{aligned} \min\quad & c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \\ \text{s. t.}\quad & x_1 + x_2 + x_3 + x_4 = 10 \\ & x \ge 0, \end{aligned} \tag{EX}$$

where $c = (c_1, c_2, c_3, c_4)$ is a cost parameter coming from the discrete uncertainty set

$$\mathcal{U} = \{(86, 12, 86, 23), (47, 33, 97, 33), (55, 94, 21, 76), (67, 40, 98, 56)\}.$$