Math. Program., Ser. A 90: 263–272 (2001)

Digital Object Identifier (DOI) 10.1007/s101070100216

Igor Averbakh

On the complexity of a class of combinatorial optimization problems with uncertainty

Received: July 1998 / Accepted: May 2000

Abstract. We consider a robust (minmax-regret) version of the problem of selecting p elements of minimum total weight out of a set of m elements with uncertainty in weights of the elements. We present a polynomial algorithm with the order of complexity $O((\min\{p, m-p\})^2 m)$ for the case where uncertainty is represented by means of interval estimates for the weights. We show that the problem is NP-hard in the case of an arbitrary finite set of possible scenarios, even if there are only two possible scenarios. This is the first known example of a robust combinatorial optimization problem that is NP-hard in the case of scenario-represented uncertainty but is polynomially solvable in the case of the interval representation of uncertainty.

Key words. polynomial algorithm – complexity – robust optimization

1. Introduction

Combinatorial optimization problems with uncertainty in input data have attracted significant research efforts because of their importance for practice. Two ways of modeling uncertainty are usually used: the stochastic approach and worst-case analysis.

In the stochastic approach, uncertainty is modelled by means of assuming some probability distribution over the space of all possible scenarios (where a scenario is a specific realization of all parameters of the problem), and the objective is to find a solution with good probabilistic performance. Models of this type are handled using stochastic programming techniques [3,5].

In the worst-case approach, the set of possible scenarios is described deterministically, and one is looking for a solution that performs reasonably well for all scenarios, i.e. that has the best "worst-case" performance. Specifically, the *minmax-regret*, or *robust* version of the worst-case approach seeks to minimize the worst-case loss in the objective function value that may occur because the solution is chosen without knowing which scenario will take place. In other words, the minmax regret approach seeks to find a solution that is ε-optimal for any possible realization of parameters, with ε as small as possible. Terms "minmax-regret" and "robust" are used interchangeably in this paper, although there are several different robustness concepts in the literature (see, e.g., [7]).

Minmax-regret combinatorial optimization problems have received increasing attention over the last decade. A comprehensive treatment of the state of art (up to 1997) in minmax-regret discrete optimization and extensive references can be found in the book [6]. However, there still are more open problems in this area than solved ones. Of

I. Averbakh: Division of Management, University of Toronto at Scarborough, 1265 Military Trail, Scarborough, Ontario M1C 1A4, Canada, e-mail: averbakh@scar.utoronto.ca

particular interest is the computational complexity of minmax-regret combinatorial optimization problems. Most of known results correspond to scenario-represented models of uncertainty (i.e. where there are finite number of possible scenarios and each one of them is given explicitly by listing the corresponding values of parameters). It has been observed that most classical polynomially solvable combinatorial optimization problems become NP-hard in minmax-regret version with scenario-represented uncertainty [6]. However, there are very few results on the complexity of minmax-regret combinatorial optimization problems with interval structure of uncertainty (where the set of possible scenarios is described by specifying an *interval of uncertainty* for each parameter and assuming that a parameter can take on any value from its interval of uncertainty). The case of interval structure of uncertainty is usually more difficult to analyze. In particular, it has not been known so far whether there are robust combinatorial optimization problems that are NP-hard in the case of scenario-represented uncertainty but are polynomially solvable in the case of interval representation of uncertainty. In this paper, we present the first example of a problem of this type. Specifically, we consider the robust version of the problem of finding the minimum weight base of a uniform matroid of rank p on a ground set of cardinality m (simply speaking, selecting a combination of p objects of minimum total weight out of a set of m objects, where there is uncertainty in weights). We show that this problem is NP-hard in the case of scenario-represented uncertainty (even if there are only two possible scenarios); however, for the case of interval representation of uncertainty we present a polynomial algorithm with complexity $O((\min\{p, (m-p)\})^2 m)$. The importance of the results is in the following:

1) They settle the complexity of this class of problems;

2) The developed polynomial algorithm can have direct practical applications, since the problem represents a class of resource allocation problems and a class of scheduling problems;

3) The ideas used for developing the polynomial algorithm may be useful for studying other robust combinatorial optimization problems with interval structure of uncertainty;

4) So far, robust combinatorial optimization problems that are NP-hard in the case of scenario-represented uncertainty have usually been conjectured intractable also in the case of interval representation. The results of this paper show that this need not to be the case, and hopefully will stimulate new efforts in trying to obtain polynomial algorithms for the large variety of robust combinatorial optimization problems whose complexity in the practically important case of interval representation of uncertainty is still open.

The paper is organized as follows. In Sect. 2, notation and problem statement are presented. In Sect. 3, we present a polynomial algorithm for the case of interval representation of uncertainty. In Sect. 4, we prove that the problem is NP-hard in the case of scenario-represented uncertainty, even if there are only two possible scenarios.

2. Notation and definitions

Let E be a finite set, $|E| = m$, and $A \subset 2^E$ be a set of *feasible subsets* of E. A vector $S = \{w_e^S, e \in E\}$ of m real numbers will be called a *scenario* and will represent an assignment of weights w_e^S to elements $e \in E$. For a scenario S and a feasible subset

$X \in A$, let us define

$$F(S, X) = \sum_{e \in X} \{w_e^S\}.$$

A generic additive combinatorial optimization problem can be formulated as follows:

Problem OPT(S). Minimize $\{F(S, X) \mid X \in A\}$.

Let $F^*(S)$ denote the optimal objective function value for Problem OPT(S). For an $X \in A$, the value $F(S, X) - F^*(S)$ is called the *regret* for the feasible set X under scenario S.

Suppose that a set SS of possible scenarios is fixed. For any $X, Y \in A$, let us define value $REGR(X, Y)$ as

$$REGR(X, Y) = \max_{S \in SS}(F(S, X) - F(S, Y)). \tag{1}$$

For an $X \in A$, let us define the worst-case regret $Z(X)$ as follows:

$$Z(X) = \max_{S \in SS} \max_{Y \in A} \{F(S, X) - F(S, Y)\}. \tag{2}$$

Alternative ways to write (2) are:

$$Z(X) = \max_{S \in SS}(F(S, X) - F^*(S)) \tag{3}$$

and

$$Z(X) = \max_{Y \in A} REGR(X, Y). \tag{4}$$

The generic robust (or minmax regret) combinatorial optimization problem corresponding to Problem OPT(S) is

Problem ROB. Find $X \in A$ that minimizes $Z(X)$.

Let Z^* denote the optimal objective function value for Problem ROB. Notice that Problem OPT(S) is a special case of Problem ROB corresponding to set SS consisting of a single scenario.

In this paper, we consider $A = \{X \subset 2^E \mid |X| = p\}$ for some fixed $p < m$, i.e., A is the set of all subsets of E with cardinality p. Then, Problem OPT(S) can be solved trivially in $O(m)$ time for any $S \in SS$ (just take p elements of E with smallest weights under scenario S). From the point of view of the underlying combinatorial structure, A is the set of bases of a uniform matroid of rank p on the ground set E [8]. From the application standpoint, Problem OPT(S) is a basic resource allocation problem [4]; it can also be interpreted as a scheduling problem (elements of E represent jobs, $-w_e^S$ is the benefit (profit) associated with processing job $e \in E$, p is the restriction (capacity) on the number of jobs that can be processed; it is required to select jobs to process so as to maximize the total benefit).

For any integers $k, s, k \le s$, notation $[k : s]$ denotes the set of all integers between k and s (including k and s).

3. The case of interval representation of uncertainty

In this section, we suppose that set SS is defined by means of specifying lower and upper bounds w_e^- and w_e^+ respectively for the weight w_e of every element $e \in E$, i.e. SS is the Cartesian product of the *intervals of uncertainty* $[w_e^-, w_e^+]$, $e \in E$. Problem ROB in this case will be referred to as Problem ROB1. Conceptually, the minmax regret approach to optimization problems with uncertainty in the case of interval representation has some connections with the tolerance approach of R.Wendell in sensitivity analysis [9], in the sense that both approaches consider simultaneous variations of all parameters of the problem rather than variations of a single parameter.

In this section, we present an algorithm for solving Problem ROB1 that has worst-case complexity $O((\min\{p, (m-p)\})^2 m)$. For simplicity of presentation, we make the following assumptions that hold throughout this section unless stated otherwise.

Assumption 1. *All $2m$ bounds w_e^-, w_e^+ are distinct (i.e. unequal) numbers.*

This is not a restrictive assumption; the general case can be handled with the same techniques using a small random perturbation of equal bounds or using some consistent rule for breaking ties (e.g., lexicographic).

Assumption 2. $p \le \frac{1}{2}m$.

Later, we show that the problem with $p > \frac{1}{2}m$ can be reduced to an equivalent problem with $p \le \frac{1}{2}m$.

For an $X \in A$, an optimal solution to the right-hand side of (3) is called a *worst-case scenario for* X; an optimal solution to the right-hand side of (4) is called a *worst-case alternative for* X.

Lemma 1. *For any $X \in A$, a worst-case scenario for X can be obtained as follows: Assign upper bounds w_e^+ as weights to all elements $e \in X$, and assign lower bounds w_e^- as weights to all elements $e \in E \setminus X$.*

Proof. For any $Y \in A$ and any $S \in SS$, value $F(S, X) - F(S, Y)$ cannot decrease if we replace S with the scenario described in Lemma 1. The statement of the lemma follows immediately.

□

The worst-case scenario for X obtained as in Lemma 1 will be denoted $S^*(X)$, i.e.

$$w_e^{S^*(X)} = \begin{cases} w_e^+, & \text{if } e \in X, \\ w_e^-, & \text{if } e \in E \setminus X, \end{cases}$$

$e \in E$. A worst-case alternative for X can be obtained by selecting p elements of E with smallest weights under the scenario $S^*(X)$; this worst-case alternative for X will be denoted $Y^*(X)$.

Let X^* be an optimal solution to Problem ROB1. Consider $S^*(X^*)$ and $Y^*(X^*)$. Let γ be the weight of the heaviest element of $Y^*(X^*)$ under the scenario $S^*(X^*)$, i.e.

$$\gamma = \max\left\{w_e^{S^*(X^*)} | e \in Y^*(X^*)\right\}. \tag{5}$$

Let e' denote the maximizer in the right-hand side of (5). Let
$B_1 = \{e \in X^* \mid w_e^+ \leq \gamma\}$,
$B_2 = \{e \in X^* \mid {w_e}^+ > \gamma\}$,
$B_3 = \{e \in E \setminus X^* \mid w_e^- \leq \gamma\}$,
$B_4 = \{e \in E \setminus X^* \mid w_e^- > \gamma\}$.

Clearly $E = B_1 \cup B_2 \cup B_3 \cup B_4$; element e' belongs to either B_1 or B_3. Notice that in the definition of sets B_1, B_2, B_3, and B_4 we use weights corresponding to scenario $S^*(X^*)$ ($w_e = w_e^+$ if $e \in X^*$ and $w_e = w_e^-$ if $e \in E \setminus X^*$). Let us define also sets
$B_2' = \{e \in B_2 \mid w_e^- > \gamma\}$,
$B_2'' = B_2 \setminus B_2'$,
$B_3' = \{e \in B_3 \mid w_e^+ \leq \gamma\}$,
$B_3'' = B_3 \setminus B_3'$.

The following three observations follow directly from the definitions and Assumption 1.

Observation 1. $X^* = B_1 \cup B_2$; $Y^*(X^*) = B_1 \cup B_3$; sets B_1, B_2, B_3, B_4 do not have common elements.

Observation 2. a) $|B_2| = p - |B_1| = |B_3|$;
b) $|B_1| \leq p$;
c) $|B_4| \geq m - 2p$.

Observation 3.

$$Z^* = Z(X^*) = \sum_{e \in B_2} w_e^+ - \sum_{e \in B_3} w_e^- . \tag{6}$$

The following three observations will be proven using interchange arguments.

Observation 4. For any $e_2 \in B_2'$ and $e_4 \in B_4$, $w_{e_4}^+ \geq w_{e_2}^+$.

Observation 5. For any $e_1 \in B_1$ and $e_3 \in B_3'$, $w_{e_1}^- \leq w_{e_3}^-$.

Observation 6. For any $e_2 \in B_2''$ and $e_3 \in B_3''$, $w_{e_2}^+ + w_{e_2}^- \leq w_{e_3}^+ + w_{e_3}^-$.

Proof of Observation 4. Suppose $e_2 \in B_2'$ and $e_4 \in B_4$ are such that $w_{e_4}^+ < w_{e_2}^+$. Then value $Z(X^*)$ can be decreased by replacing e_2 with e_4 in X^*, since according to (6) the change of $Z(X^*)$ will be $w_{e_4}^+ - w_{e_2}^+ < 0$ (notice that $Y^*(X^*) \setminus X^* = B_3$ does not change after such replacement), which contradicts the optimality of X^* for Problem ROB1.

□

Proof of Observation 5. Suppose $e_1 \in B_1$ and $e_3 \in B_3'$ are such that $w_{e_1}^- > w_{e_3}^-$. Then value $Z(X^*)$ can be decreased by replacing e_1 with e_3 in X^*, since according to (6) the change of $Z(X^*)$ will be $w_{e_3}^- - w_{e_1}^- < 0$ (notice that $X^* \setminus Y^*(X^*) = B_2$ does not change after such replacement), which contradicts the optimality of X^* for Problem ROB1.

□

Proof of Observation 6. Suppose $e_2 \in B_2''$ and $e_3 \in B_3''$ are such that $w_{e_2}^+ + w_{e_2}^- > w_{e_3}^+ + w_{e_3}^-$. Then value $Z(X^*)$ can be decreased by replacing e_2 with e_3 in X^*, since according to (6) the change of $Z(X^*)$ will be $w_3^+ - w_2^+ - (w_2^- - w_3^-) = w_3^+ + w_3^- - w_2^+ - w_2^- < 0$, which contradicts the optimality of X^* for Problem ROB1.

□

Suppose that values γ and $r = |B_1|$ are known. Then $|B_2| = |B_3| = p - r$, $|B_4| = m - 2p + r$, and X^* can be obtained as follows:

1. Obtain set $C_1 = \{e \in E \mid w_e^+ \le \gamma\}$ (Clearly $C_1 = B_1 \cup B_3'$).

2. Choose r elements e of C_1 with smallest w_e^-; this is set B_1 (according to Observation 5). The remaining $|C_1| - r$ elements of C_1 form set B_3'.

3. Obtain set $C_2 = \{e \in E \mid w_e^- > \gamma\}$ (Clearly $C_2 = B_4 \cup B_2'$)

4. Choose $m - 2p + r$ elements e of C_2 with largest w_e^+; this is set B_4 (according to Observation 4). The remaining $|C_2| - (m - 2p + r)$ elements of C_2 form set B_2'.

5. Elements of $E \setminus C_1 \setminus C_2$ form $B_2'' \cup B_3''$ (sets B_1, B_2', B_3', B_4 have already been obtained). We know that $|B_2''| = |B_2| - |B_2'|$ and $|B_2| = |B_3| = p - r$. Select $p - r - |B_2'|$ elements of $E \setminus C_1 \setminus C_2$ with smallest values of $w_e^+ + w_e^-$ (breaking ties arbitrarily); this is B_2'' (using Observation 6.) Now, $X^* = B_1 \cup B_2' \cup B_2''$.

Given values r and γ, the above procedure can be implemented in $O(m)$ time (note that selecting k smallest numbers out of a set of m numbers takes $O(m)$ time [1]). Of course we do not know values r and γ in advance; but we can try all combinations of candidate values for r and γ, applying a procedure similar to the one above for each combination and then choosing the best of the obtained candidate solutions (for any $X \in A$, value $Z(X)$ can be obtained in $O(m)$ time using Lemma 1). There are $p+1$ candidate values for r (integers 0,1,...,p). We know that $\gamma = w_e^+$ or $\gamma = w_e^-$ for some $e \in E$; therefore, there are $2m$ candidate values for γ. However, it is sufficient to consider a smaller set of candidate values for γ, using the following observations.

Observation 7. $\gamma = w_e^+$ for some $e \in B_1$ or $\gamma = w_e^-$ for some $e \in B_3$.

Observation 8. a) For any $e_3 \in B_3$, and any $e_4 \in B_4$, $w_{e_3}^- < w_{e_4}^-$;
b) For any $e_1 \in B_1$ and any $e_4 \in B_4$, $w_{e_1}^+ < w_{e_4}^+$.

Let U_1 be the set of $2p$ smallest numbers out of $\{w_e^+ \mid e \in E\}$; let U_2 be the set of $2p$ smallest numbers out of $\{w_e^- \mid e \in E\}$; and let $U = U_1 \cup U_2$. Then Observation 8 along with Observation 2,c imply that for any $e \in B_3$, $w_e^- \in U_2$, and for any $e \in B_1$, $w_e^+ \in U_1$. With Observation 7, we obtain the following:

Observation 9. $\gamma \in U$.

Thus, it is sufficient to consider the set U of cardinality $4p$ as the set of candidate values for γ.

This discussion justifies the algorithm for solving Problem ROB1 described below. Speaking informally, the algorithm tries all values $\hat{r} \in [0 : p]$ and $\hat{\gamma} \in U$ as candidates for the true values r and γ, keeping record of the best tentative solution obtained so far (denoted X') and the corresponding objective function value (denoted Z'). The algorithm uses Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X') that "checks" whether the input values of $\hat{r}$ and $\hat{\gamma}$

can be used to improve upon the recorded solution. Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X') is guaranteed to obtain an optimal solution X^* when the input values of $\hat{r}$ and $\hat{\gamma}$ are equal to the true values of r and γ.

Algorithm 1.
Preprocessing: Obtain the set U described above ($|U| = 4p$). Set $X' =$ (any member of A) (X' will denote the best solution obtained so far), $Z' = Z(X')$ (value $Z(X')$ can be obtained in $O(m)$ time using Lemma 1).

Begin
  For every $\hat{r} \in [0 : p]$ do begin
    For every $\hat{\gamma} \in U$ do begin
      Apply Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X');
    End;
  End;
  Set $X^* = X'$, $Z^* = Z'$;
  Output X^*, Z^*;
End

Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X')
Input: $\hat{r}$, $\hat{\gamma}$, Z', X'.
Output: Z', X'
Begin

1. Obtain set $C_1(\hat{\gamma}) = \{e \in E \mid w_e^+ \leq \hat{\gamma}\}$. If $\hat{r} > |C_1(\hat{\gamma})|$, STOP (in this case $\hat{\gamma}$ and $\hat{r}$ cannot be true values for γ and r).
2. Choose $\hat{r}$ elements e of $C_1(\hat{\gamma})$ with smallest w_e^-; denote the set of the chosen elements as $B_1(\hat{\gamma}, \hat{r})$, and the set of remaining $|C_1(\hat{\gamma})| - \hat{r}$ elements of $C_1(\hat{\gamma})$ as $B_3'(\hat{\gamma}, \hat{r})$.
3. Obtain set $C_2(\hat{\gamma}) = \{e \in E \mid w_e^- > \hat{\gamma}\}$. If $m - 2p + \hat{r} > |C_2(\hat{\gamma})|$, STOP (in this case $\hat{\gamma}$ and $\hat{r}$ cannot be true values for γ and r).
4. Choose $m - 2p + \hat{r}$ elements e of $C_2(\hat{\gamma})$ with largest w_e^+; denote the set of these elements as $B_4(\hat{\gamma}, \hat{r})$, and the set of remaining $|C_2(\hat{\gamma})| - (m - 2p + \hat{r})$ elements of $C_2(\hat{\gamma})$ as $B_2'(\hat{\gamma}, \hat{r})$. If $|B_2'(\hat{\gamma}, \hat{r})| > p - \hat{r}$, or if $p - \hat{r} - |B_2'(\hat{\gamma}, \hat{r})| > |E \setminus C_1(\hat{\gamma}) \setminus C_2(\hat{\gamma})|$, STOP (in this case $\hat{\gamma}$ and $\hat{r}$ cannot be true values for γ and r).
5. Select $p - \hat{r} - |B_2'(\hat{\gamma}, \hat{r})|$ elements of $E \setminus C_1(\hat{\gamma}) \setminus C_2(\hat{\gamma})$ with smallest values of $w_e^+ + w_e^-$ (breaking ties arbitrarily); denote the set of these elements as $B_2''(\hat{\gamma}, \hat{r})$.
6. Set $X''(\hat{\gamma}, \hat{r}) = B_1(\hat{\gamma}, \hat{r}) \cup B_2'(\hat{\gamma}, \hat{r}) \cup B_2''(\hat{\gamma}, \hat{r})$ ($|X''(\hat{\gamma}, \hat{r})| = p$, since $|B_1(\hat{\gamma}, \hat{r})| = \hat{r}$ and $|B_2'(\hat{\gamma}, \hat{r})| + |B_2''(\hat{\gamma}, \hat{r})| = p - \hat{r}$).
7. Compute $Z(X''(\hat{\gamma}, \hat{r}))$ (this can be done in $O(m)$ time using Lemma 1.)
8. If $Z(X''(\hat{\gamma}, \hat{r})) < Z'$, update $Z' \leftarrow Z(X''(\hat{\gamma}, \hat{r}))$, $X' \leftarrow X''(\hat{\gamma}, \hat{r})$.

End

Theorem 1. *Algorithm 1 correctly solves Problem ROB1 in $O(p^2 m)$ time.*

Proof. Since $|U| = 4p$, the algorithm applies Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X') $O(p^2)$ times. The complexity of Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X') is $O(m)$; the complexity bound $O(p^2 m)$ for Algorithm 1 follows.

For any combination of $\hat{\gamma} \in U$ and $\hat{r} \in [0 : p]$, Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X') either does not change value Z' and the tentative solution X', or finds a solution $X''(\hat{\gamma}, \hat{r})$ that is better than the current tentative solution X'; in this case, Z' and X' are updated. At every iteration of the algorithm, $Z' \geq Z^*$. Since $\gamma \in U$ and $r \in [0 : p]$, Algorithm 1 will eventually apply Procedure CHECK($\hat{r}$, $\hat{\gamma}$, Z', X') with $\hat{r} = r$, $\hat{\gamma} = \gamma$; previous discussion implies that at this iteration an optimal solution X^* will be obtained. Therefore, Algorithm 1 ends with $X' = X^*$ and $Z' = Z^*$.

□

In this section, we assumed $p \leq \frac{1}{2}m$, and Algorithm 1 was developed using this assumption. Suppose now that Assumption 2 does not hold and $p > \frac{1}{2}m$. One way to handle this case is to include new "dummy" elements in E with sufficiently large weight bounds w_e^-, w_e^+ (so that these "dummy" elements can never be in an optimal solution), until $|E|$ becomes larger than $2p$, and then apply Algorithm 1; the complexity of Algorithm 1 in this case is $O(m^3)$. Below we describe a better approach that reduces Problem ROB1 to a problem of the same type (in some sense, dual to the original Problem ROB1) for which Assumption 2 holds. The complexity of Algorithm 1 in this case will be $O((m-p)^2 m)$ which is better than $O(m^3)$ if $(m-p) = o(m)$.

Let $\tilde{A} = \{\tilde{X} \subset 2^E || \tilde{X}| = m - p\}$, i.e. $\tilde{A}$ is the set of all subsets of E with cardinality $m - p$. Problem OPT(S) where set A is replaced with set $\tilde{A}$ will be called **Problem OPTD(S)**; the optimal objective function value for Problem OPTD(S) will be denoted $\tilde{F}^*(S)$. Let $\widetilde{SS}$ be the Cartesian product of intervals $[-w_e^+, -w_e^-]$, $e \in E$. For a scenario $S = \{w_e^S, e \in E\}$, let $-S$ denote the scenario $\{-w_e^S, e \in E\}$, i.e. $w_e^{-S} = -w_e^S$ for any $e \in E$. For any $\tilde{X} \in \tilde{A}$, let $\tilde{F}(S, \tilde{X}) = \sum_{e \in \tilde{X}} \{w_e^S\}$, and let $\tilde{Z}(\tilde{X}) = \max_{S \in SS} (\tilde{F}(-S, \tilde{X}) - \tilde{F}^*(-S)) = \max_{\tilde{S} \in \widetilde{SS}} (\tilde{F}(\tilde{S}, \tilde{X}) - \tilde{F}^*(\tilde{S}))$.

Observation 10. For any scenario S, $F^*(S) - \tilde{F}^*(-S) = \sum_{e \in E} w_e^S$.

Problem ROB1 where set A is replaced with $\tilde{A}$ and set SS is replaced with $\widetilde{SS}$ will be called **Problem ROB1D**. That is, Problem ROB1D is to find $\tilde{X} \in \tilde{A}$ that minimizes $\tilde{Z}(\tilde{X})$.

Theorem 2. *If $\tilde{X}^* \in \tilde{A}$ is an optimal solution to Problem ROB1D, then $X^* = E \setminus \tilde{X}^*$ is an optimal solution to Problem ROB1.*

Proof. For any $X \in A$, consider $\tilde{X} = E \setminus X$. Clearly $\tilde{X} \in \tilde{A}$. Notice that for any $S \in SS$, $\sum_{e \in X} w_e^S - \sum_{e \in \tilde{X}} w_e^{-S} = \sum_{e \in E} w_e^S$. Taking into account Observation 10, we have that

$$F(S, X) - F^*(S) = \tilde{F}(-S, \tilde{X}) - \tilde{F}^*(-S),$$

and therefore $Z(X) = \tilde{Z}(\tilde{X})$. The statement of the theorem follows immediately.

□

Algorithm 1 and the bound in Theorem 1 were developed assuming $p \leq \frac{1}{2}m$. If $p > \frac{1}{2}m$, we can apply Algorithm 1 to Problem ROB1D (with complexity bound $O((m-p)^2 m)$); according to Theorem 2, this will also give us an optimal solution to Problem ROB1. Thus, we have

Theorem 3. *Problem ROB1 can be solved in $O((\min\{p, m-p\})^2 m)$ time.*

4. The case of scenario-represented uncertainty

In this section, we suppose that the set SS of possible scenarios is finite, and each scenario S from SS is given explicitly (by specifying w_e^S, $e \in E$). Problem ROB in this case is referred to as Problem ROB2. We show in this section that Problem ROB2 is NP-hard even if there are only two possible scenarios ($|SS| = 2$).

Theorem 4. *Problem ROB2 is NP-hard even if SS contains only two scenarios.*

Proof. We use a reduction from the following decision problem that is known to be NP-complete ([2]).

Problem PARTITION. Given is a set E', $|E'| = 2n$. Each element $e \in E'$ has a positive integer weight w'_e. Is there a subset $X' \subset E'$ such that $|X'| = n$ and $\sum_{e \in X'} w'_e = K$, where $K = \frac{1}{2} \sum_{e \in E'} w'_e$?

Reduction. For an instance of Problem PARTITION (defined by set E' and positive weights w'_e, $e \in E'$) we define the corresponding instance of Problem ROB2 as follows. Let r_1 be the total weight of $n - 1$ heaviest elements of E', and r_2 be the total weight of $n - 1$ lightest elements of E'. Set $E = E' \cup \{b_1, b_2\}$, $p = n$, $SS = \{S_1, S_2\}$, $w_e^{S_1} = w'_e$ for any $e \in E'$, $w_e^{S_2} = -w'_e$ for any $e \in E'$, $w_{b_1}^{S_1} = -r_2$, $w_{b_1}^{S_2} = 10K$, $w_{b_2}^{S_2} = r_1 - 2K$, $w_{b_2}^{S_1} = 10K$. Now, we claim that the obtained instance of Problem ROB2 has the optimal objective value Z^* equal to K if and only if the answer to the original instance of PARTITION was "yes"; otherwise, $Z^* > K$. Then, the NP-hardness of Problem ROB2 follows from the NP-completeness of PARTITION.

To prove the claim, we notice that $F^*(S_1) = 0$, $F^*(S_2) = -2K$, and neither b_1 nor b_2 can belong to any optimal solution to the obtained instance of Problem ROB2 (weights $w_{b_1}^{S_2}$ and $w_{b_2}^{S_1}$ have been chosen sufficiently large to prevent such a possibility: For any $X \in A$, if X contains b_1 or b_2, then $Z(X) \geq 6K$; and if X contains neither b_1 nor b_2, then $Z(X) \leq 4K$). So, any optimal solution to the instance of Problem ROB2 contains only elements of E'.

Now, for any $X \in A$ such that X contains neither b_1 nor b_2, $F(S_2, X) - F^*(S_2) = 2K - (F(S_1, X) - F^*(S_1))$ (since $F(S_1, X) = -F(S_2, X)$, $F^*(S_1) = 0$, $F^*(S_2) = -2K$). Therefore, $Z(X) = \max\,\{F(S_1, X) - F^*(S_1), F(S_2, X) - F^*(S_2)\} = \max\,\{F(S_1, X), 2K - F(S_1, X)\} = \max\,\left\{\sum_{e \in X} w'_e, 2K - \sum_{e \in X} w'_e\right\}$. Thus, $Z(X) \geq K$ for any $X \in A$. If $Z^* = K$, then there is $X \in A$ such that $Z(X) = K$, X is an optimal solution to Problem ROB2 and, as observed above, X does not contain b_1 or b_2; therefore, $X \subset E'$ and $\sum_{e \in X} w'_e = K$, i.e. the answer to the original instance of PARTITION is "yes". On the other hand, if $Z^* > K$, then there is no $X \in A$ such that $F(S_1, X) = K$, and the answer to the original instance of PARTITION is "no". The theorem is proven.

□

Remark. It is natural to ask the question about the complexity of Problem ROB in the case where set SS is represented by a general system of linear constraints on weights, i.e. where SS is a general polytope in R^m. Problem ROB in this case will be referred to as Problem ROB3. Clearly Problem ROB1 is a special case of Problem ROB3. Using Theorem 4, it is not difficult to show that Problem ROB3 is NP-hard, by means of a polynomial reduction from Problem ROB2 with two scenarios to Problem ROB3.

The reduction is obtained as follows: the polytope corresponding to set SS for Problem ROB3 is just the convex hull of the two scenarios of Problem ROB2 in R^m (i.e. the line segment with endpoints at the scenarios). It is easy to see that the instance of Problem ROB3 obtained by the reduction is equivalent to the original instance of Problem ROB2. Thus, we have the following

Corollary 1. *Problem ROB3 is NP-hard.*

5. Conclusion

In the paper, we considered the minmax regret version of the problem of selecting p objects of minimum total weight out of a set of m objects, assuming uncertainty in the weights of objects. For the problem with interval structure of uncertainty, where weights can take on any values from the corresponding intervals of uncertainty (and, therefore, the set of scenarios is a rectangular box in the space of weights), we presented a polynomial algorithm with the order of complexity $O((\min\{p, m-p\})^2 m)$. We proved that the problem with scenario-represented structure of uncertainty is NP-hard even in the case of just two scenarios. As a corollary to this result, we obtained that the problem with the set of scenarios represented by an arbitrary polytope is NP-hard (thus, polynomial solvability of the interval data case is due to the special structure of the set of scenarios which is a rectangular box).

To our knowledge, this is the first known example of a minmax regret combinatorial optimization problem that is polynomially solvable in the case of interval representation of uncertainty while being NP-hard in the case of scenario-represented uncertainty. It shows that there is no direct relationship between the complexity of the case of scenario representation of uncertainty and the complexity of the interval-data case. Hopefully, this will stimulate new efforts in trying to obtain polynomial algorithms for numerous interval data minmax regret combinatorial optimization problems whose complexity is still open but which are known to be NP-hard in the case of scenario-represented uncertainty.

References

1. Aho, A.V., Hopcroft, J.E., Ullman, J.D. (1974): The Design and Analysis of Computer Algorithms. Addison-Wesley, Reading, MA
2. Garey, M., Johnson, D. (1979): Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, New York
3. Ermolyev, Ju.M. (1976): Methods of Stochastic Programming. Nauka, Moscow
4. Ibaraki, T., Katoh, N. (1988): Resource Allocation Problems. MIT Press, Cambridge, Massachusetts
5. Kall, P., Wallace, S. (1994): Stochastic Programming. Wiley, New York
6. Kouvelis, P., Yu, G. (1997): Robust Discrete Optimization and Its Applications. Kluwer Academic Publ., Boston
7. Mulvey, J.M., Vanderbei, R.J., Zenios, S.A. (1995): Robust Optimization of Large Scale Systems. Oper. Res. **43**, 264–281
8. Welsh, D.J. (1976): Matroid Theory. Academic Press, London
9. Wendell, R.E. (1984): Using Bounds on the Data in Linear Programming: the Tolerance Approach to Sensitivity Analysis. Math. Program. **28**, 304–322