Operations Research Letters 36 (2008) 618–622

Contents lists available at ScienceDirect

Operations Research Letters

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

Polymatroids and mean-risk minimization in discrete optimization

Alper Atamtürk*, Vishnu Narayanan

Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720-1777, USA

ARTICLE INFO

Article history:
Received 24 January 2008
Accepted 28 April 2008
Available online 11 June 2008

Keywords:
Conic integer programming
Value-at-risk
Ellipsoidal uncertainty sets
Submodular functions

ABSTRACT

We study discrete optimization problems with a submodular mean-risk minimization objective. For 0–1 problems a linear characterization of the convex lower envelope is given. For mixed 0–1 problems we derive an exponential class of conic quadratic valid inequalities. We report computational experiments on risk-averse capital budgeting problems with uncertain returns.

1. Introduction

In financial markets, high levels of risk are associated with big returns as well as big losses; on the other hand, with lower levels of risk, the potential for return or loss is small. Risk management is fundamentally concerned with finding an optimal trade-off between risk and return matching an investor's risk tolerance. Although studied mostly in a financial context, managing risk is relevant in any area with a significant source of uncertainty.

The mean-risk optimization is well-studied for problems with a convex feasible set [18,22]. However, this is not the case in the discrete setting, even though, portfolios are often restricted to discrete choices in practice. In this paper, we study mean-risk minimization for problems with discrete decision variables. In particular, we consider

$$\min \left\{ f(x) = \sum_i \mu_i x_i + \Omega \sqrt{\sum_i \sigma_i^2 x_i^2} : x \in \mathcal{F} \right\}, \tag{1}$$

where $\mathcal{F} \subseteq \{0,1\}^n \times [0,1]^m$. Problem (1) often arises when minimizing a stochastic objective over a discrete feasible set. For example, if μ_i and σ_i^2 are the mean and variance of independent normally distributed random variables ℓ_i (loss or negative return on investment i), for $0 < \epsilon < 0.5$, by setting $\Omega = -\Phi^{-1}(\epsilon)$, where Φ is the c.d.f. of the standard normal distribution, (1) is equivalent to the value-at-risk (VaR) minimization problem (e.g. Birge and Louveaux [14]):

$$\zeta(\epsilon) := \min \left\{ z : \textbf{Prob}\left(\sum_i \ell_i x_i > z \right) \le \epsilon, x \in \mathcal{F} \right\}. \tag{2}$$

Here $\mathcal{F}$ denotes a constrained set of possible investments, which may include discrete choices. So, the mean-risk objective of (1) models the trade-off between long-run average and short-run risk with the parameter Ω measuring the investor's risk tolerance. We refer the reader to Ahmed [1] and references therein for stochastic optimization with more general mean-risk objectives.

If ℓ_i has an unknown distribution with partial information, then a *robust version* (a la Ben-Tal and Nemirovski [5,8], El Ghaoui et al. [17]) of VaR minimization (2) can be written as problem (1) by appropriately choosing Ω. For example, if only the first two moments μ_i, σ_i^2 of ℓ_i are known, then a robust version based on extremal probability distributions with such moments can be written as (1) by letting $\Omega = \sqrt{(1-\epsilon)/\epsilon}$ (Bertsimas and Popescu [10], El Ghaoui et al. [16]).

We mention a few earlier uses of the model (1) in a discrete setting. Ishii et al. [20] consider a stochastic spanning tree problem, where edge lengths are i.i.d. Normal(μ_i, σ_i^2) and formulate it as (1) with $\mathcal{F}$ denoting the set of spanning trees of a graph. Ozsen et al. [25] and Shen et al. [27] define *risk pooling* models of the form (1) for integrated warehouse location and inventory. In these models the objective captures fixed and transportation costs as well as the cost of maintaining safety stock for uncertain retailer demand during the delivery lead time, and $\mathcal{F}$ denotes the set of feasible warehouse locations and corresponding retailer assignments. Vielma et al. [28] solve discrete portfolio optimization problems with a risk constraint with a general branch-and-bound algorithm based on the linear approximation of the conic quadratic cone due to Ben-Tal and Nemirovski [7].

Problem (1) is also referred to as *robust discrete optimization*. For the 0–1 case, i.e., when $m = 0$, Bertsimas and Sim [13] describe an

* Corresponding author.
E-mail addresses: atamturk@berkeley.edu (A. Atamtürk), vishnu@ieor.berkeley.edu (V. Narayanan).

0167-6377/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2008.04.006

algorithm converging to a locally optimal solution. Bertsimas and Sim [11,12] give a linear approximation for the mean-risk objective of problem (1). Atamtürk [2] gives several strong alternative formulations for this linear model.

When restricted to only binary variables, the objective of function f of (1) is submodular. In Section 2 we show a connection between the convex lower envelope of submodular f and extended polymatroids. Then, in Section 3 we give a complete linear description of the convex lower envelope of the mean risk function of (1). In Section 4 we consider the generalization to mixed 0–1 problems and derive valid conic quadratic constraints for the corresponding convex lower envelope. In Section 5 we discuss computational experiments on using these results to solve a mean-risk capital budgeting problem with discrete choices. Finally, we finish with a few closing remarks in Section 6 indicating that our results indeed hold for a generalization of (1).

2. Preliminaries

In this section we consider the minimization of a set function and give a simple result on its convex lower envelope, which we use for the mean-risk minimization objective of (1). Let $N := \{1, \ldots, n\}$ be a finite set and $f : 2^N \to \mathbb{R}$ be a set function on N. We are interested in minimizing f, i.e.,

$$\min_{S \subseteq N} f(S). \tag{3}$$

Without loss of generality, we may assume that $f(\emptyset) = 0$, since otherwise we can solve the equivalent minimization problem for $f' := f - f(\emptyset)$, i.e., the shifted function f' with $f'(S) = f(S) - f(\emptyset)$ for all $S \subseteq N$.

Throughout the paper, by abusing notation, we will refer to a set function also as $f(x)$, where $x \in \{0, 1\}^n$ is the indicator vector for subsets of N. Let χ_S denote the indicator vector of a set S and S_x denote the support set of a vector x. Now using this notation, let us we rewrite the minimization problem (3) as $\min \left\{z : (x, z) \in \mathcal{Q}_f\right\}$, where

$$\mathcal{Q}_f := \text{conv}\{(x, z) \in \{0, 1\}^n \times \mathbb{R} : f(x) \le z\}.$$

It is clear that $\mathcal{Q}_f$ is the convex lower envelope of f. Because it is the convex hull of disjunction of 2^n polyhedra (for each assignment of x), $\mathcal{Q}_f$ is a polyhedron as well. For a set function f on N satisfying $f(\emptyset) = 0$, let

$$EP_f := \{\pi \in \mathbb{R}^N : \pi(S) \le f(S) \text{ for all } S \subseteq N\},$$

where $\pi(S)$ denotes $\sum_{i \in S} \pi_i,\ S \subseteq N$. The next simple proposition shows a polarity relationship between EP_f and a subset of the valid inequalities for $\mathcal{Q}_f$.

Proposition 1. *Inequality $\pi x \le z$ is valid for $\mathcal{Q}_f$ if and only if $\pi \in EP_f$.*

Proof. For $\pi \in EP_f$, we have $\pi x = \pi(S_x) \le f(S_x) \le z$. Conversely, if $\pi \notin EP_f$, then $\pi(S) > f(S)$ for some $S \subseteq N$; but then for $z = f(S)$, $\pi(S) = \pi \chi_S > z$, contradicting the validity of $\pi x \le z$. □

If the function f is submodular, then inequalities of Proposition 1 have a nice characterization as shown by Edmonds [15].

Definition 1. A set function $f : 2^N \to \mathbb{R}$ is *submodular* if

$$f(S) + f(T) \ge f(S \cup T) + f(S \cap T) \quad \text{for all } S, T \subseteq N.$$

For a survey on submodular function minimization we refer the reader to [19,21]. If f is a submodular function, EP_f is called the *extended polymatroid* associated with f [26]. For an extended polymatroid, we call inequalities $\pi x \le z$ with $\pi \in EP_f$ as the *extended polymatroid inequalities* of $\mathcal{Q}_f$. Edmonds [15] showed that π is an extreme point of extended polymatroid EP_f if and only if $\pi_i = f(S_{(i)}) - f(S_{(i-1)})$, where $S_{(i)} = \{(1), (2), \ldots, (i)\}$, $1 \le i \le n$ for some permutation $((1), (2), \ldots, (n))$ of N. We will refer to inequalities $\pi x \le z$ defined by the extreme points of the extended polymatroid EP_f as the *extremal extended polymatroid inequalities*.

The separation problem for extended polymatroid inequalities is optimization of a linear objective over EP_f, which can be solved by the greedy algorithm [15]: Given $\bar{x} \in \mathbb{R}^n_+$ and $\bar{z} \in \mathbb{R}$, checking whether $(\bar{x}, \bar{z})$ violates an extended polymatroid inequality is equivalent to solving the problem $\zeta := \max\left\{\pi\bar{x} : \pi \in EP_f\right\}$. For a nonincreasing order $\bar{x}_{(1)} \ge \bar{x}_{(2)} \ge \cdots \ge \bar{x}_{(n)}$, let $S_{(i)} = \{(1), (2), \ldots, (i)\}$ and $\bar{\pi}_{(i)} = f(S_{(i)}) - f(S_{(i-1)})$ for $1 \le i \le n$. Then, the point $(\bar{x}, \bar{z})$ is violated by an extended polymatroid inequality if and only if $\zeta = \bar{\pi}\bar{x} > \bar{z}$.

Remark 1. Note that if $f(\emptyset) \ne 0$, the extended polymatroid inequalities for $\mathcal{Q}_f$ take the form

$$\pi x \le z - f(\emptyset), \quad \pi \in EP_{f'},$$

where $f' := f - f(\emptyset)$.

3. 0–1 optimization

In this section we consider minimizing the mean-risk objective of (1) with only binary variables:

$$\min\left\{g(x) := ax + \Omega\sqrt{cx + \sigma^2} : x \in \{0, 1\}^n\right\}, \tag{4}$$

where $\Omega, \sigma, c \ge 0$. Notice that we replaced x_i^2 in (1) with x_i in (4) as they are equivalent for $x_i \in \{0, 1\}$. Note, however, whereas the objective of (1) is a convex function, the objective of (4) is concave over $\mathbb{R}^n_+$.

If $\Omega = 0$, the problem is trivial; otherwise, by scaling the objective, we assume that $\Omega = 1$. Also, without loss of generality, we assume that $c_i > 0$ for all i, because if $c_i = 0$, then x_i can be set to either 0 or 1, depending on the sign of a_i. We index the variables so that $\frac{a_1}{c_1} \le \cdots \le \frac{a_n}{c_n}$ (breaking the ties arbitrarily) and let $S_i := \{1, 2, \ldots, i\}$ for $i = 1, 2, \ldots, n$.

Proposition 2. *The set of all optimal solutions to (4) is some collection $\mathcal{S}$ of nested sets $S_{i_1} \subset S_{i_2} \subset \cdots \subset S_{i_k}$, $1 \le k \le n$.*

Proof. This proposition is a slight strengthening of Theorem 4.2 in [27] and concerns *all* optimal solutions to a given problem rather than *some*. For completeness, we repeat the argument here, by also observing the strict concavity of the square root function. For every optimal solution S and $i < j, j \in S$ implies $i \in S$. To see this, suppose $i \notin S$ and consider the objective values z^*, z', and z'' corresponding to $S, S \cup i$ and $S \setminus j$, respectively. Then, for some $\delta \ge 0$

$$\frac{z' - z^*}{c_i} = \frac{a_i}{c_i} + \frac{\sqrt{\delta + c_i + c_j} - \sqrt{\delta + c_j}}{c_i} < \frac{a_j}{c_j} + \frac{\sqrt{\delta + c_j} - \sqrt{\delta}}{c_j} = \frac{z^* - z''}{c_j} \le 0,$$

where the strict inequality holds by strict the concavity of the square root function. Thus, $z' < z^*$, contradicting optimality of S. □

Remark 2. Proposition 2 implies that for any choice of data there is at most one optimal solution of a given cardinality. Hence, there are at most n distinct optimal solutions, which are nested. This is because while indexing variables, if $\frac{a_i}{c_i} = \frac{a_j}{c_j}$, then for any optimal solution S we must have either $\{i, j\} \subseteq S$ or $\{i, j\} \cap S = \emptyset$.

Proposition 3. *The mean-risk function g is submodular; and so is $g - \sigma$.*

Proof. From the concavity of the square root function and $c \geq \mathbf{0}$, we see that the difference function $\rho_i(S) := g(S \cup i) - g(S)$, $S \subset N$ and $i \in N \setminus S$, is nonincreasing, i.e., $\rho_i(S) \geq \rho_i(T)$ for all $S \subseteq T \subset N$ and all $i \in N \setminus T$, which is equivalent to Definition 1 of submodularity [23,24]. □

Theorem 1. *$\mathcal{Q}_g$ is described completely by extended polymatroid inequalities*

$$\pi x \leq z - \sigma, \ \pi \in EP_{g-\sigma} \tag{5}$$

and bound inequalities $0 \leq x \leq 1$.

Proof. Consider the set of optimal solutions for $\min\{\alpha x + \beta z : (x, z) \in \mathcal{Q}_g\}$ with $(\alpha, \beta) \neq (\mathbf{0}, 0)$. We may assume that $\beta \geq 0$, since otherwise the problem is unbounded. If $\beta = 0$, then there exists $i \in N$ with $\alpha_i \neq 0$. If $\alpha_i > 0$, $x_i = 0$ for all optimal solutions. If $\alpha_i < 0$, $x_i = 1$ for all optimal solutions. Then, we may assume, if necessary by scaling, that $\beta = 1$. Now, observe that the optimization problem $\min\left\{\alpha x + z : (x, z) \in \mathcal{Q}_g\right\}$ is equivalent to the minimization of the submodular function $h(S) := \sum_{i \in S} \alpha_i + g(S)$, whose optimal solutions $\mathcal{S}$ are nested in a nondecreasing order of $(\alpha_i + a_i)/c_i$ by Proposition 2. Then, the extremal extended polymatroid inequality $\pi x \leq z - \sigma$, $\pi \in EP_{h-\sigma}$ with this order satisfies $\pi(S) = h(S) - \sigma$ for all $S \in \mathcal{S}$. Now let $\pi' = \pi - \alpha$. Because $\pi'(S) = \pi(S) - \sum_{i \in S} \alpha_i$, $\pi' \in EP_{g-\sigma}$ and $\pi'(S) = g(S) - \sigma$ for all $S \in \mathcal{S}$. Hence each optimal solution $(x, z) = (\chi_S, g(S))$ is on the face of $\mathcal{Q}_g$ defined by $\pi' x \leq z - \sigma$. □

4. Mixed 0–1 optimization

In this section we consider the mean-risk minimization objective with binary as well as continuous variables

$$\min\left\{ax + by + \Omega\left(cx + \sum_{i=1}^{m} d_i y_i^2\right)^{1/2} : (x, y) \in \mathcal{F}\right\}, \tag{6}$$

where $\mathcal{F} = \{0, 1\}^n \times [0, 1]^m$ and $\Omega, c, d > \mathbf{0}$. Indeed, if necessary, by scaling the objective we assume that $\Omega = 1$. Now let us state problem (6) as $\min\{z : (x, y, z) \in \mathcal{R}_h\}$, where

$$\mathcal{R}_h := \operatorname{conv}\{(x, y, z) \in \{0, 1\}^n \times [0, 1]^m \times \mathbb{R} : h(x, y) \leq z\},$$

and

$$h(x, y) := ax + by + \left(cx + \sum_{i=1}^{m} d_i y_i^2\right)^{1/2}.$$

Unlike the 0–1 case in the previous section, $\mathcal{R}_h$ is not a polyhedral set. The function h is convex in y and concave in x over the domain. Note that for any fixed value of x, problem (6) reduces to a conic quadratic optimization problem (Ben-Tal and Nemirovski [6]) that can be solved easily. Also from Proposition 2, we know that for any fixed value of y, the optimal binary solution has a nested structure, *which is independent of the value of y* (the ordering of the variables in Proposition 2 is not a function of σ^2). Therefore, problem (6) can be solved in polynomial time, by fixing binary variables in the nested order S_i, $i = 1, \ldots, n$ one at a time and then solving the remaining conic quadratic optimization problem in polynomial time.

In the following, we derive conic quadratic valid inequalities of the form

$$\pi x + by + \left(\sum_{i \in T} d_i y_i^2\right)^{1/2} \leq z, \quad T \subseteq \{1, 2, \ldots, m\} \tag{7}$$

for $\mathcal{R}_h$ by using the results from the 0–1 case. Toward this end, for $T \subseteq \{1, 2, \ldots, m\}$ let us derive a set function $f_T : 2^N \to \mathbb{R}$ as $f_T(x) := h(x, \chi_T)$. Clearly, f_T is submodular on N. Let $\sigma_T := f_T(\mathbf{0}) = b(T) + d(T)^{1/2}$.

Proposition 4. *Inequality (7) is valid for $\mathcal{R}_h$ if and only if $\pi \in EP_{f_T - \sigma_T}$.*

Proof. Suppose $\pi \in EP_{f_T - \sigma_T}$. For $(x, y) \in \{0, 1\}^n \times [0, 1]^m$

$$\begin{aligned}
\pi x + by + \left(\sum_{i \in T} d_i y_i^2\right)^{1/2} &\leq f_T(x) - \sigma_T + by + \left(\sum_{i \in T} d_i y_i^2\right)^{1/2} \\
&= ax + by + (cx + d(T))^{1/2} - d(T)^{1/2} + \left(\sum_{i \in T} d_i y_i^2\right)^{1/2} \\
&\leq ax + by + \left(cx + \sum_{i=1}^{m} d_i y_i^2\right)^{1/2} \leq z,
\end{aligned}$$

where the last line follows from

$$\begin{aligned}
&(\delta + d(T))^{1/2} - d(T)^{1/2} \\
&= \min_{\mathbf{0} \leq y \leq \mathbf{1}} \left(\delta + \sum_{i \in T} d_i y_i^2\right)^{1/2} - \left(\sum_{i \in T} d_i y_i^2\right)^{1/2}
\end{aligned}$$

for any $\delta \geq 0$ by the concavity of the square root function and $d \geq \mathbf{0}$. Therefore, $(x, y, z) \in \mathcal{R}_h$.

Conversely, if $\pi \notin EP_{f_T - \sigma_T}$, then for some $\bar{x} \in \{0, 1\}^n$ we have $\pi \bar{x} > f_T(\bar{x}) - \sigma_T$. But, then for $(\bar{x}, \sum_{i \in T} e_i, \bar{z})$ with $\bar{z} = a\bar{x} + b(T) + (c\bar{x} + d(T))^{1/2}$, inequality (7) is violated. □

Note that inequalities (7) reduce to their linear counterparts (5) when $y = \chi_T$. Because these inequalities have the same structure as (5), for fixed T the separation problem for them is also a polymatroid optimization problem.

5. Computational experience

In this section we present our computational experiments on using the inequalities developed for solving a risk-averse capital budgeting problem

$$\begin{aligned}
\zeta = \max &\left\{\mu x - \Omega(\epsilon)\sqrt{\sum_i \sigma_i^2 x_i^2}\right. \\
&\left. : \sum_i a_i x_i \leq b, x \in \{0, 1\}^n \times [0, 1]^m\right\}.
\end{aligned} \tag{8}$$

As usual, μ_i and σ_i^2 denote the mean and variance of uncertain return r_i on investment i. We assume that the returns are independent. Then with $\Omega(\epsilon) = \sqrt{(1 - \epsilon)/\epsilon}$ for small $\epsilon > 0$, the return of the portfolio is at least ζ with a probability greater than $1 - \epsilon$ for an optimal solution to (8) [10,16].

For the computational experiments, we used the MIP solver of CPLEX 11.0 that solves SOCP relaxations at the nodes of the branch-and-bound tree, with CPLEX heuristics turned off. In Table 1 we report on the experiments with pure 0–1 decisions for varying problem sizes and probabilistic guarantees. For each combination, five random instances are generated with μ_i and a_i from uniform $[0, 100]$, and σ_i from uniform $[0, \mu_i]$. The budget b is set to $0.5 \cdot \sum_i a_i$. In the table we compare the integrality gap of the SOCP relaxation, number of nodes explored and the CPU time (in seconds) with and without adding extended polymatroid cuts (5). The cuts are generated only at the root node with the greedy algorithm as explained in Section 2. We observe that, in general,

Table 1
Pure 0–1 problems

n	ϵ	CPLEX			CPLEX + cuts (5)			
		% gap	Nodes	Time	Cuts	% gap	Nodes	Time
	.10	5.67	250	1	4	0.73	47	0
	.05	12.43	585	2	4	1.06	115	1
25	.03	24.43	2 345	6	7	2.67	125	1
	.02	32.62	843	3	5	1.88	52	0
	.01	43.21	315	1	5	0.65	17	0
	.10	2.88	1 129	34	3	0.30	193	3
	.05	6.46	4 228	33	4	0.33	199	2
50	.03	10.24	29 957	214	5	0.31	134	2
	.02	14.67	98 530	646	7	0.66	911	18
	.01	26.05	205 290	1076	8	1.52	16 439	79
	.10	0.96	3 025	30	3	0.10	308	7
	.05	2.35	13 375	145	4	0.09	192	3
100	.03	4.96	76 809	911	5	0.14	475	25
	.02	7.96	182 603	1873[a]	5	0.13	978	181
	.01	15.81	204 104	1884[a]	10	0.26	2 904	831

[a] Instances could not be solved in 30 mins.

Table 2
Mixed 0–1 problems ($n = 100$)

m	ϵ	CPLEX			CPLEX + cuts (7)			
		% gap	Nodes	Time	Cuts	% gap	Nodes	Time
	.10	0.81	556	7	2	0.23	92	1
	.05	2.09	10 792	139	3	0.54	567	8
5	.03	4.40	55 452	788	4	0.62	9 124	156
	.02	6.97	149 853	1859	5	0.90	66 819	1158
	.01	13.68	167 627	1871[a]	7	2.12	107 531	1849[a]
	.10	0.74	569	8	2	0.19	124	2
	.05	1.91	9116	138	3	0.29	985	16
10	.03	4.04	42 451	709	4	0.67	12 950	206
	.02	6.38	99 521	1511	5	1.13	64 403	1296
	.01	12.15	139 219	1857[a]	6	2.86	80 997	1838[a]
	.10	0.63	571	10	3	0.09	109	2
	.05	1.63	8 974	153	3	0.26	895	22
20	.03	3.41	33 259	665	4	0.66	7 478	278
	.02	5.29	72 535	1386	6	1.27	24 248	1110
	.01	9.97	114 365	1852[a]	8	3.52	40 816	1830[a]

[a] Instances could not be solved in 30 mins.

as the probabilistic guarantee increases, so does the integrality gap of the original SOCP formulation. This can be explained by the increasing weight of the risk term in the objective, which typically leads to a high number of fractional variables in the continuous relaxation. Note that the number of branch-and-bound nodes and the CPU time increases with the problem size.

We note that for $n = 100$ without cuts, none of the five instances for $\epsilon = 0.02, 0.01$ could be solved to optimality within the time limit of 30 mins. The average optimality gap at termination for these unsolved instances are 3.30% and 10.79%, respectively. On the other hand, when cuts are added, the integrality gap at the root node of the tree reduces to less than one percent for almost all instances, which translates to a significant reduction in the number of nodes as well in the CPU time. Recall that by Theorem 1 the remaining small gap at the root node and branching are due to the budget constraint.

In Table 2 we report the results for mixed 0–1 problems with 100 binary variables and varying number (m) of continuous variables. In these experiments, we have used the nonlinear cuts (7). As before, the cuts were generated only at the root node. We used a simple heuristic that picks $T \subseteq \{1, \ldots, m\}$ for the nonlinear cut (7) to add. For a given $(\bar{x}, \bar{y})$, an index i is included T with probability 0 if $\bar{y}_i \leq 0.33$, with probability 0.5 if $0.33 < \bar{y}_i < 0.66$, and with probability 1 if $\bar{y}_i \geq 0.66$. Once T is fixed, the coefficients π for the binary variables are computed using the separation algorithm in Section 2. In the table we see a significant reduction in the integrality gap, the number of nodes explored as well as the total CPU time with the addition of the cuts. None of the instances with $\epsilon = 0.01$ could be solved to optimality for $m = 5, 10, 15$ within the time limit of 30 minutes. The average optimality gap left at termination for these unsolved instances are 8.76, 7.39, 5.84%, respectively, when no cuts are added compared with 1.59, 3.60, 2.71%, respectively, when the cuts are added. Hence, for these instances, although an optimal solution is not found either, the optimality gap is reduced significantly when the cuts are used.

6. Final remarks

We have shown that the convex lower envelope of the mean-risk function of (1) can be described with simple bounds and extended polymatroid inequalities. Our results in Section 3 extend to

$$g(x) = ax + h(c_o + cx)$$

for strictly concave h and $c_o \geq 0$, $c > \mathbf{0}$. Therefore, the results in Section 4 for the mixed 0–1 case hold, for instance, for the robust convex objective defined with the L_p norm (Bertsimas et al. [9])

$$h(x) = \mu x + \Omega \left(\sum_i (\sigma_i x_i)^p \right)^{1/p}$$

with $\Omega > 0$, $\sigma > \mathbf{0}$ and $p > 1$. To which other functions the results in the paper extend is an interesting question.

The inequalities may be useful computationally, even in cases where they give a partial characterization. Our computational study for testing the impact of the inequalities as cuts for minimizing value at risk in capital budgeting problems shows promise. For the mixed 0–1 case additional research, perhaps, based on the application of general techniques for conic mixed-integer programming, such as conic mixed-integer rounding [3] and conic lifting [4], is warranted.

Acknowledgements

We are grateful to Shabbir Ahmed for thoughtful comments on an earlier version of the paper and discussions on related topics. This research is funded, in part, by the National Science Foundation Grant 0700203. The first author is grateful for the hospitality of the Georgia Institute of Technology, where part of this research was conducted.

References

[1] S. Ahmed, Convexity and decomposition of mean-risk stochastic programs, Mathematical Programming 106 (2006) 433–446.
[2] A. Atamtürk, Strong formulations of robust mixed 0–1 programming, Mathematical Programming 108 (2006) 235–250.
[3] A. Atamtürk, V. Narayanan, Cuts for conic mixed integer programming, in: M. Fischetti, D.P. Williamson (Eds.), in: Proceedings of the 12th International IPCO Conference, 2007, pp. 16–29.
[4] A. Atamtürk, V. Narayanan, Lifting for conic mixed-integer programming, Technical Report BCOL.07.04, IEOR, University of California-Berkeley, 2007.
[5] A. Ben-Tal, A. Nemirovski, Robust convex optimization, Mathematics of Operations Research 23 (1998) 769–805.
[6] A. Ben-Tal, A. Nemirovski, Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications, in: MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.
[7] A. Ben-Tal, A. Nemirovski, On polyhedral approximations of the second-order cone, Mathematics of Operations Research 26 (2001) 193–205.
[8] A. Ben-Tal, A. Nemirovski, Robust optimization–methodology and applications, Mathematical Programming 92 (2002) 453–480.
[9] D. Bertsimas, D. Pachamanova, M. Sim, Robust linear optimization under general norms, Operations Research Letters 32 (2004) 510–516.
[10] D. Bertsimas, I. Popescu, Optimal inequalities in probability theory: A convex optimization approach, SIAM Journal on Optimization 15 (2005) 78–804.
[11] D. Bertsimas, M. Sim, Robust discrete optimization and network flows, Mathematical Programming 98 (2003) 49–71.
[12] D. Bertsimas, M. Sim, The price of robustness, Operations Research 52 (2004) 35–53.

[13] D. Bertsimas, M. Sim, Robust discrete optimization under ellipsoidal uncertainty sets, Manuscript, MIT, 2004.
[14] J.R. Birge, F. Louveaux, Introduction to Stochastic Programming, Springer, Berlin, 1997.
[15] J. Edmonds, Submodular functions, matroids and certain polyhedra, in: R. Guy (Ed.), in: Combinatorial Structres and their Applications, vol. 11, Gordon and Breach, New York, NY, 1971, pp. 69–87.
[16] L. El Ghaoui, M. Oks, F. Oustry, Worst-case value-at-risk and robust portfolio optimization: A conic programming approach, Operations Research 51 (2003) 543–556.
[17] L. El Ghaoui, F. Oustry, H. Lebret, Robust solutions to uncertain semidefinite programs, SIAM Journal on Optimization 9 (1998) 33–52.
[18] H. Föllmer, A. Schied, Stochastic Finance: An Introduction in Discrete Time, Walter de Gruyter, Berlin, 2004.
[19] S. Fujishige, Submodular Functions and Optimization, 2nd ed., in: Annals of Discrete Mathematics, vol. 58, Elsevier, 2005.
[20] H. Ishii, S. Shiode, T. Nishida, Stochastic spanning tree problem, Discrete Applied Mathematics 3 (1981) 263–273.
[21] S. Iwata, Submodular function minimization, Mathematical Programming 112 (2008) 45–64.
[22] H. Markowitz, Mean-Variance Analysis in Portfolio Choice and Capital Markets, Basil Blackwell, New York, 1987.
[23] G.L. Nemhauser, L.A. Wolsey, Integer and Combinatorial Optimization, John Wiley and Sons, New York, 1988.
[24] G.L. Nemhauser, L.A. Wolsey, M.L. Fisher, An analysis of approximations for maximizing submodular set functions-i, Mathematical Programming 14 (1978) 265–294.
[25] L. Ozsen, M.S. Daskin, C.R. Coullard, Capacitated warehouse location model with risk pooling, Naval Research Logistics 55 (2008) 295–312.
[26] A. Schrijver, Combinatorial Optimization: Polyhedra and Efficiency, Springer Verlag, Berlin, 2003.
[27] Z.-J.M. Shen, C. Coullard, M.S. Daskin, A joint location-inventory model, Transportation Science 37 (2003) 40–55.
[28] J.P. Vielma, S. Ahmed, G.L. Nemhauser, A lifted linear programming branch-and-bound algorithm for mixed integer conic quadratic programs, Manuscript, Georgia Institute of Technology, 2007.