Operations Research Letters 43 (2015) 445–449

Contents lists available at ScienceDirect

Operations Research Letters

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

Robust binary optimization using a safe tractable approximation

Ruken Düzgün [a], Aurélie Thiele [b,*]

[a] Marriott International, Bethesda MD, USA
[b] Lehigh University, Bethlehem PA, USA

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form
10 June 2015
Accepted 11 June 2015
Available online 18 June 2015

Keywords:
Robust optimization
Binary problems
Chance constraints
Bernstein approximation

ABSTRACT

We present a robust optimization approach to 0–1 linear programming with uncertain objective coefficients based on a safe tractable approximation of chance constraints, when only the first two moments and the support of the random parameters are known. We obtain nonlinear problems with only one additional (continuous) variable. Our robust optimization problem can be interpreted as a nominal problem with modified coefficients. We compare our approach with Bertsimas and Sim (2003). In numerical experiments, we obtain solutions of similar quality in faster time.

1. Introduction

We consider binary optimization problems with uncertain objective coefficients and investigate the models that arise from enforcing probabilistic constraints on the objective in the context of robust optimization. Robust optimization is a worst-case optimization approach where the worst case is computed over a given uncertainty set that the unknown parameters belong to, centered around their nominal values and of size reflecting the degree of the decision maker's aversion to ambiguity. The reader is referred to Bertsimas et al. [4] for a review of robust optimization up to 2011 and Gabrel et al. [12] for a review of recent advances in robust optimization. Providing an intuitive interpretation of uncertainty sets has always been of importance to operations researchers: for instance, Bertsimas and Sim [9] connect the choice of a key parameter in their approach, called the budget of uncertainty, with a probability of constraint violation. More recently, Ben-Tal et al. [3] describe a process where a "safe tractable approximation" of probabilistic constraints leads to a robust optimization problem where the uncertainty set is determined by the chosen approximation and the probability level. Safe tractable approximations, the most famous of which is the Bernstein approximation, are motivated by the fact that incorporating a chance constraint to a problem creates significant computational difficulties if the random variables do not obey a jointly Gaussian distribution. In Ben-Tal et al. [3], the probabilistic constraint is replaced by a more tractable constraint that, when satisfied, guarantees that the original constraint is satisfied too.

Our goal in the present paper is to investigate the theoretical and algorithmic insights we gain from the approach presented in Ben-Tal et al. [3], in the special case where decision variables are binary, and to compare this approach with that obtained in Bertsimas and Sim [9]. Hence, we will use the same description of uncertainty based on range forecasts for each of the uncertainty parameters and a budget of uncertainty. It is worth pointing out, however, that other descriptions of uncertainty have been proposed since [9]. The construction of uncertainty sets based on available data is investigated for instance in Bertsimas et al. [6], who build their method on statistical hypothesis tests and derive finite-sample probabilistic guarantees for the optimal solutions they obtain. In the context of mixed integer optimization, Bertsimas and Dunning [5] present an adaptive partition approach based on Voronoi diagrams to Multistage Adaptive Mixed Integer Optimization, which allows them to gain insights into the regions of the uncertainty set restricting the objective function values.

Distributionally robust optimization, where the uncertainty is on the probability distributions obeyed by the random parameters rather than on the values taken by uncertain parameters themselves, has been the focus of extensive research interest, for instance in Wiesemann et al. [19], who introduce standardized ambiguity sets containing all distributions with prescribed conic representable confidence sets and provide conditions under which distributionally robust optimization problems based on their ambiguity sets are computationally tractable. Delage and Ye [11]

* Corresponding author.
E-mail addresses: rud207@lehigh.edu (R. Düzgün), aut204@lehigh.edu (A. Thiele).

http://dx.doi.org/10.1016/j.orl.2015.06.006
0167-6377/

combine the streams of distributionally robust optimization and data-driven optimization with a focus on moment uncertainty. Their model incorporates uncertainty in both the distribution form (such as Gaussian or exponential) and the first two moments (mean and covariance matrix), and can be solved efficiently for a wide range of cost functions. Distributionally robust optimization and its tractable approximations are also studied in Goh and Sim [13] in the context of a linear programming problem with uncertainties that has expectations in the objective and the constraints. Their framework leads to an approximate solution that is distributionally robust and more flexible than linear rules. Uncertainty sets in distributionally robust optimization have been studied using Kullback–Leibler divergences, which originated in information theory and measures the distance between two distributions (Hu and Hong [15]) and, more generally, Phi-divergences, of which Kullback–Leibler are a special case (Love and Bayraksan [17]). Phi-divergences representing uncertainty in optimization problems affected by uncertain probabilities are also investigated in Ben-Tal et al. [2]; the authors show that the robust counterpart of a linear optimization problem with phi-divergence uncertainty is tractable for most choices of phi typically considered in the literature. Additional references for distributionally robust optimization are provided in [12]. Robust integer programming, which is at the core of the present paper, has received less attention in the literature after [9]. Wagner [18] provides a robust formulation of a stochastic 0–1 linear programming problem for which only the first k moments of the random parameters are known. Hanasusanto et al. [14] focus on two-stage problems with integer recourse and approximate two-stage robust binary problems with K-adaptability problems, where the decision maker designs in advance (here and now) K second-stage policies and implements the best of the K policies once uncertain parameters have been observed.

We make the following contributions to the literature. (i) We provide robust formulations when we only know the first two moments and the support of the distributions of the uncertain parameters. (ii) We show that the safe tractable approximation (Bernstein approximation) in our setting can be interpreted as a deterministic problem with modified cost coefficients that only depend on problem data and one extra coefficient. (iii) We compare our approach in numerical experiments with that in Bertsimas and Sim [9] for the same problem setting but for a different modeling of uncertainty in robust optimization and argue that, while solution quality is comparable, the solution times in our approach are substantially smaller.

2. The safe tractable approximation

We consider the following problem with uncertain objective coefficients.

$$\begin{aligned} \max \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & x \in \mathcal{X} \subseteq \{0,1\}^n. \end{aligned} \tag{1}$$

Because the vector $\mathbf{c}$ is not known precisely, our goal here will be to maximize the greatest parameter A such that:

$$\mathbb{P}\left(\mathbf{c}'\mathbf{x} < A\right) \leq \epsilon, \tag{2}$$

for $\epsilon > 0$ given (small). We first assume that the random coefficients are independent, and will relax this assumption in Section 4. When distributions are continuous, A can be interpreted as the ϵ-quantile of $\mathbf{c}'\mathbf{x}$.

Specifically, the stochastic problem we consider is:

$$\begin{aligned} \max \quad & \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-VaR}_\epsilon(\mathbf{c}'\mathbf{x}) \\ \text{s.t.} \quad & x \in \mathcal{X} \subseteq \{0,1\}^n, \end{aligned} \tag{3}$$

where $\mathbb{P}\text{-VaR}_\epsilon(\tilde{\xi}) = \sup_{A\in\mathbb{R}}\{A : \mathbb{P}(\tilde{\xi} < A) \leq \epsilon\}$ denotes the ϵ-Value-at-Risk of a random variable $\tilde{\xi}$ that obeys the probability distribution $\mathbb{P}$, and $\mathcal{P}$ denotes the ambiguity set for the probability distribution, i.e., the distributions deemed compatible with the decision maker's information about the random objective coefficients.

We are interested in deriving a deterministic tractable counterpart to our problem when only a limited amount of information is known: the mean, variance and support of each uncertain parameter. Knowledge of the first two moments is a common assumption in distributionally robust optimization (see Gabrel et al. [12] and the references therein), while knowledge of the support is the foundation of the polyhedral uncertainty sets in Bertsimas and Sim [10].

Lemma 2.1. *If $\mathbb{E}[\exp\{-\theta c_i\}]$ can be computed efficiently for all i and any $\theta > 0$, a safe tractable approximation to Problem (3) is:*

$$\begin{aligned} \max_{\theta,x} \quad & \frac{\ln\epsilon}{\theta} - \frac{1}{\theta}\sum_{i=1}^{n} \ln \mathbb{E}[\exp\{-\theta c_i\}]\cdot x_i \\ s.t. \quad & x \in \mathcal{X} \subseteq \{0,1\}^n,\ \theta > 0. \end{aligned} \tag{4}$$

Proof. From Ben-Tal et al. [3], Eq. (2) can be written as, with $\theta > 0$:

$$\mathbb{P}\left(-\theta\sum_{i=1}^{n} c_i x_i > -\theta A\right) = \mathbb{P}\left(\exp\left\{-\theta\sum_{i=1}^{n} c_i x_i\right\} > \exp\{-\theta A\}\right) \leq \epsilon.$$

Since the exponential function is nonnegative and nondecreasing, we invoke Markov's Inequality and inject that the coefficients are independent:

$$\mathbb{P}\left(\mathbf{c}'\mathbf{x} < A\right) \leq \frac{\mathbb{E}\left[\exp\left\{-\theta\sum_{i=1}^{n} c_i x_i\right\}\right]}{\exp\{-\theta A\}} = \frac{\prod_{i=1}^{n}\mathbb{E}[\exp\{-\theta c_i x_i\}]}{\exp\{-\theta A\}}. \tag{5}$$

The safe tractable approximation replaces $\mathbb{P}\left(\mathbf{c}'\mathbf{x} < A\right) \leq \epsilon$ with $\frac{\prod_{i=1}^{n}\mathbb{E}[\exp\{-\theta c_i x_i\}]}{\exp\{-\theta A\}} \leq \epsilon$, thus guaranteeing that the original probabilistic constraint is satisfied. Taking the logarithm of the left and right-hand sides leads to the greatest possible value of A being $\frac{\ln\epsilon}{\theta} - \frac{1}{\theta}\sum_{i=1}^{n}\ln\mathbb{E}[\exp\{-\theta c_i x_i\}]$. We conclude by using that the x_i are binary, so $\ln\mathbb{E}[\exp\{-\theta c_i x_i\}] = \ln\mathbb{E}[\exp\{-\theta c_i\}]\cdot x_i\ \forall i$. □

However, in our problem setup, the expected values of the random parameters $\exp\{-\theta c_i\},\ i = 1,\ldots,n$ are not known exactly; instead, only the first two moments and the support of the random parameters are known. Therefore, we seek tight upper bounds of these expected values by adapting the linear semi-infinite optimization approach of Bertsimas and Popescu [7,8].

Lemma 2.2. *Consider a random parameter c. From a probabilistic perspective, let μ be its (known) mean and σ its (known) standard deviation. From a robust optimization perspective, let $\bar{c}$ be its nominal value and $\hat{c}$ be the half-width of its range forecast or confidence interval, i.e., let $[\bar{c}-\hat{c},\ \bar{c}+\hat{c}]$ be the support of the uncertain parameter. We assume $\bar{c} = \mu$. Finally, let m be a positive number such that $\hat{c} = m\,\sigma$ for $m \geq 1$ and let π be the set of such possible distributions. Then,*

$$\max_{f\in\pi}\mathbb{E}_f[\exp\{-\theta c\}] = \frac{\exp\{-\theta\bar{c}\}}{m^2+1} \cdot\left(\exp\{\theta\hat{c}\} + m^2\cdot\exp\left\{-\frac{\theta\hat{c}}{m^2}\right\}\right). \tag{6}$$