Systems & Control Letters 94 (2016) 63–69

Contents lists available at ScienceDirect

Systems & Control Letters

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

On the computational complexity and generalization properties of multi-stage and stage-wise coupled scenario programs

Nikolaos Kariotoglou [a,*], Kostas Margellos [b], John Lygeros [a]

[a] *Automatic Control Laboratory, Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland*

[b] *Control Group, Department of Engineering Science, University of Oxford, United Kingdom*

ARTICLE INFO

Article history:
Received 10 October 2015
Received in revised form
5 April 2016
Accepted 11 May 2016
Available online 11 June 2016

Keywords:
Scenario approach
Randomized optimization
Uncertain systems
Approximate dynamic programming

ABSTRACT

We discuss the computational complexity and feasibility properties of scenario sampling techniques for uncertain optimization programs. We propose an alternative way of dealing with a special class of stage-wise coupled programs and compare it with existing methods in the literature in terms of feasibility and computational complexity. We identify trade-offs between different methods depending on the problem structure and the desired probability of constraint satisfaction. To illustrate our results, an example from the area of approximate dynamic programming is considered.

1. Introduction

One way of dealing with data uncertainty in robust optimization is to allow the optimal decision to violate problem constraints on a set of pre-specified measure. The authors in [1,2] provide explicit solutions to such problems under assumptions on the probability distribution of the uncertainty. To avoid such assumptions one can make use of uncertainty samples and construct decisions that only satisfy system constraints on the sampled uncertainty instances. The *scenario approach* [3,4] can be used to provide feasibility generalization statements for convex optimization problems, i.e., how likely it is for a sample-based decision to satisfy the problem constraints for a new realization of the uncertainty that was not included in the samples. Beyond feasibility guarantees, [5–7] provide bounds on the amount of constraint violation and are concerned with probabilistic performance issues.

We focus on scenario based convex optimization problems using the scenario approach [3,4,8,9]. We illustrate that the same guarantees on the feasibility of a scenario based solution may be obtained by formulating alternative scenario programs, each with a potentially different number of decision variables and constraints and hence different computational complexity. We argue that in the case of multi-stage scenario programs, stage-wise coupled via the constraint functions, it is often challenging to decide which algorithm to use and illustrate how different choices give rise to a significant trade-off in the total computation time. Motivated by such cases, we provide a framework to compare approaches in terms of computational complexity, while sharing the same joint constraint feasibility properties. In this context, our contributions are: (1) We show how the scenario approach paradigm can be deployed in stage-wise coupled programs and analyze the feasibility properties of the associated solutions (Sections 3.3 and 3.4). (2) We illustrate how the violation and confidence parameters can be treated as additional degrees of freedom and be selected by means of a convex program in view of reducing the computational complexity (Section 3.5). This is fundamentally different to the existing literature where violation and confidence levels are typically considered as fixed parameters when computing sample size bounds. (3) We compare alternatives with respect to computational complexity and identify underlying trade-offs (Section 4). (4) We demonstrate the results on an approximate dynamic programming (ADP) algorithm developed for reachability problems (Section 5). Applications are not limited to this algorithm since most ADP approaches based on [10] result in a sequence of coupled scenario programs.

Section 2 presents the general problem under consideration and Section 3 the different scenario based alternatives and their properties. In Section 4 we discuss the trade-off between feasibility and computational complexity of each alternative while Section 5 illustrates our results with a numerical example in ADP.

* Corresponding author.
E-mail addresses: nkarioto@ethz.ch (N. Kariotoglou), kostas.margellos@eng.ox.ac.uk (K. Margellos), lygeros@ethz.ch (J. Lygeros).

http://dx.doi.org/10.1016/j.sysconle.2016.05.009

Let $\mathbb{R}$, $\mathbb{N}$, $\mathbb{N}_+$ denote the real, natural and positive natural numbers. Uncertainty samples are extracted from a possibly unknown set Δ according to a possibly unknown probability measure $\mathbb{P}$. $\mathbb{P}^S$ with $S \in \mathbb{N}_+$ denotes the corresponding product measure. We use i.i.d for identically and independently distributed. Operator $|\cdot|$ denotes the cardinality of its argument, $\dim(A)$ the dimension of a linear space A, and $x \models y$ that x satisfies statement y.

2. Programs with multiple robust constraints

Consider a compact convex set $\mathcal{X} \subseteq \mathbb{R}^d$, a possibly unbounded uncertainty set $\Delta \subseteq \mathbb{R}^w$, a convex cost function $f : \mathcal{X} \to \mathbb{R}$ and a set of $M \in \mathbb{N}_+$ convex constraint functions $g_i : \mathcal{X} \times \Delta \to \mathbb{R}$, $i = 1, \dots, M$. We deal with robust convex optimization problems (RCP) of the form:

$$\text{RCP}: \begin{cases} \min_{x\in\mathcal{X}} & f(x) \\ \text{s.t} & g_i(x,\delta) \le 0, \quad \forall \delta \in \Delta,\ \forall i \in \{1,\dots,M\}. \end{cases} \tag{1}$$

The set Δ may be infinite and possibly unbounded, rendering (1) a convex, semi-infinite optimization program. A common approach to approximate the solution is to impose the constraints on a finite number of uncertainty instances. Consider $S \in \mathbb{N}_+$ i.i.d samples $\{\delta^j\}_{j=1}^S$ extracted from Δ according to some underlying probability distribution, and a collection $\{\Delta_i\}_{i=1}^M$ of M subsets of $\{\delta^j\}_{j=1}^S$ such that for each $\delta \in \{\delta^j\}_{j=1}^S$ there exists i so that $\delta \in \Delta_i$, i.e., the sets may be overlapping but each δ belongs in at least one of them. The interpretation is that for each $i = 1, \dots, M$, the corresponding constraint $g_i(x, \delta)$ should be satisfied for all $\delta \in \Delta_i$, but not necessarily for all $\delta \in \Delta$. Problem (1) is then approximated by a scenario convex optimization program (SCP) of the form:

$$\text{SCP}[\Delta_1,\dots,\Delta_M]: \begin{cases} \min_{x\in\mathcal{X}} & f(x) \\ \text{s.t} & g_i(x,\delta) \le 0, \quad \forall \delta \in \Delta_i, \\ & \forall i \in \{1,\dots,M\} \end{cases} \tag{2}$$

and can be solved to optimality by various solvers. We impose the following assumption on $\text{SCP}[\Delta_1, \dots, \Delta_M]$:

Assumption 1. For any set $\{\delta^j\}_{j=1}^S$ and collection of subsets $\{\Delta_i\}_{i=1}^M$ with $S, M \in \mathbb{N}_+$, $\text{SCP}[\Delta_1, \dots, \Delta_M]$ is feasible, its feasibility region has a non-empty interior and its minimizer $x^*[\Delta_1, \dots, \Delta_M] : \Delta^S \to \mathcal{X}$ is unique.

We refer to [3,11] for details on how Assumption 1 can be relaxed. Measurability of $x^*[\Delta_1, \dots, \Delta_M]$ is assumed as needed [7,12]. Same as the standard literature on the scenario approach [3,4] we focus on the feasibility properties of x^* as a function of the algorithm used to construct it; performance issues are discussed in [6,7].

3. Feasibility of scenario convex programs

We introduce four different approaches to formulate the SCP: the standard scenario approach, the multi-stage scenario approach, the stage-wise coupled scenario approach using the same samples at every step and the stage-wise coupled scenario approach using different samples at every step. Due to differences in the generation of samples, each approach provides different design choices. In particular, whenever the constraints in the SCP are sampled separately, additional degrees of freedom are introduced, allowing to choose different feasibility properties for each constraint. We compare all approaches on the same metric of jointly satisfying all of the constraints in (2), and exploit their structure to reduce computational complexity.

3.1. The standard scenario approach

Let $\bar{\Delta} = \{\delta^j\}_{j=1}^S$ and assume that $\Delta_1 = \cdots = \Delta_M = \bar{\Delta}$, i.e., enforce each constraint on all elements in $\bar{\Delta}$. Let d be the dimension of the decision space $\mathcal{X}$ and denote by $\text{SCP}[\bar{\Delta}]$, $x^*[\bar{\Delta}]$ the resulting instance of $\text{SCP}[\Delta_1, \dots, \Delta_M]$ and its minimizer, respectively. According to [4, Theorem 2.4], one can choose violation and confidence levels $\varepsilon, \beta \in (0, 1)$, sample

$$S \ge S(\varepsilon, \beta, d) \tag{3}$$

with $S(\varepsilon,\beta,d) := \min\left\{N \in \mathbb{N} \middle| \sum_{i=0}^{d-1} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} \le \beta\right\}$ points from the constraint set of (1) according to $\mathbb{P}$ and formulate $\text{SCP}[\bar{\Delta}]$ where $\Delta_1 = \cdots = \Delta_M = \bar{\Delta}$ are constructed using the extracted samples. Under Assumption 1, the minimizer of the resulting problem, $x^*[\bar{\Delta}]$, satisfies

$$\text{CCP}_\varepsilon : \mathbb{P}[\exists i \in \{1,\dots,M\},\ g_i(x^*[\bar{\Delta}],\delta) > 0] \le \varepsilon \tag{4}$$

with confidence (measured with respect to $\mathbb{P}^S$) at least $1 - \beta$. The final statement can be compactly written as $\mathbb{P}^S[x^*[\bar{\Delta}] \models \text{CCP}_\varepsilon] \ge 1 - \beta$. The computational complexity associated with constructing $x^*[\bar{\Delta}]$, along with its feasibility properties depend on the choice of ε, β and the number of decision variables d that implicitly affect the number of constraints (inspect (3)). Note that the result remains unaffected if d in (3) is replaced by any upper bound on the number of the so-called support constraints (see [3] for a precise definition) other than the dimension of the decision space. Refinements along this direction are discussed in [8,13,14] where the authors present a tighter bound, defined as the constraint support rank.

3.2. The multi-stage scenario approach

We impose additional structure on the RCP by assuming that for any $\delta \in \Delta$ and each $i = 1, \dots, M$, the constraint function $g_i(\cdot, \delta)$ does not necessarily depend on all decision variables. The set-up is then similar to the structure considered in [8], where the authors studied optimization programs with multiple chance constraints. For each $i = 1, \dots, M$, let $\mathcal{X}_i \subseteq \mathcal{X}$ denote the domain of each $g_i(\cdot, \delta)$ and $d_i = \dim(\mathcal{X}_i)$, where $\dim(\mathcal{X}_i)$ denotes the dimension of the smallest subspace of $\mathbb{R}^d$ containing $\mathcal{X}_i$. We further assume that $d_i < d$ for at least one $i = 1, \dots, M$ to exclude the case where all constraint functions depend on all decision variables; if this is not the case the subsequent analysis reduces to the standard scenario approach of Section 3.1. It was shown in [8, Theorem 4.1] that one can choose different violation and confidence levels $\varepsilon_i, \beta_i \in (0, 1)$ for each $i = 1, \dots, M$, extract

$$S_i \ge \sum_{i=1}^{M} S(\varepsilon_i, \beta_i, d_i) \tag{5}$$

with

$$S(\varepsilon_i,\beta_i,d_i) := \min\left\{N \in \mathbb{N} \middle| \sum_{j=0}^{d_i-1} \binom{N}{j} \varepsilon_i^j (1-\varepsilon_i)^{N-j} \le \beta_i\right\}$$

samples i.i.d from Δ according to a probability measure $\mathbb{P}$, construct $\{\Delta_i\}_{i=1}^M$ as in Section 2 with $|\Delta_i| = S_i$ and formulate $\text{SCP}[\Delta_1, \dots, \Delta_M]$. Under Assumption 1, it holds that for each $i = 1, \dots, M$, the minimizer $x^*[\Delta_1, \dots, \Delta_M]$ of $\text{SCP}[\Delta_1, \dots, \Delta_M]$ satisfies the chance constraint,

$$\text{CCP}_{\varepsilon_i} : \mathbb{P}[g_i\left(x^*[\Delta_1,\dots,\Delta_M],\delta\right) > 0] \le \varepsilon_i, \tag{6}$$

with confidence (measured with respect to $\mathbb{P}^{S_i}$) at least $1 - \beta_i$. As with the standard scenario approach in Section 3.1, each d_i can be replaced by a tighter upper bound on the support