

Efficient Algorithms And Representations For Chance-constrained Mixed Constraint Programming

Cheng Fang

Massachusetts Institute of Technology
 Computer Science and Artificial Intelligence Laboratory
 32 Vassar Street, Cambridge MA, 02139
 cfang@mit.edu

Abstract

Resistance to autonomous systems comes in part from the perceived unreliability of the systems. Concerns can be addressed by guarantees of the probability of success. This is achieved in chance-constrained constraint programming (CC-CP) by imposing constraints required for success, and providing upper-bounds on the probability of violating constraints. This extended abstract reports on novel uncertainty representations to address problems prevalent in current methods.

1 Introduction

In decision making under uncertainty, a subset of decisions are under the control of the agent, and the rest are assigned by the environment. Chance-constrained constraint programming (CC-CP) allows the specification of a set of constraints over *controllable variables* assigned by the agent, and *uncontrollable variables* assigned by the environment, as well as a probabilistic bound on satisfying subsets of the constraints. Current weakness in CC-CP are as follows:

- Current chance-constrained problems are typically cast as nonlinear optimisation problems, and processed with generalised solvers (eg IPOPT [Wächter and Biegler, 2006], SNOPT [Gill *et al.*, 2005]). Feasible problems are typically small. Faster solution methods are required to extend the chance-constrained approach to real-world problems.
- Current chance-constrained techniques rely on reasonable probability distributions to represent uncertainty. When distributions are fitted from historical data, for example with kernel density estimation (KDE), there are concerns with over-fitting. The alternative, prevalent in robust optimisation, is to construct a unique robust set from data [Bertsimas *et al.*, 2017]. Such techniques do not take the objective function or constraints into consideration when constructing the robust sets, leading to over-conservatism.

2 Problem Formulation And Approach

The problem to be investigated is a mixed constraint system, with variables with integer and real domains, and linear and nonlinear constraints over the variables. Formally:

Problem 1. (Data-driven cc-MINLP)

Let $\langle \Omega, \mathcal{F}, \mathbb{P} \rangle$ be a probability space and $\mathbf{u} : \Omega \rightarrow \mathbb{R}^{N_U}$ be vector of random variables, where $\mathbf{u}^{(j)}$ is the j -th random variable. Given a set of samples $\{\hat{\mathbf{u}}_1^{(j)}, \hat{\mathbf{u}}_2^{(j)}, \dots, \hat{\mathbf{u}}_{M_j}^{(j)}\}$ for each random variable $\mathbf{u}^{(j)}$,

$$\begin{aligned} & \min_{\mathbf{x}} \mathbb{E} [(c(\mathbf{x}, \mathbf{u}(\omega)))] \\ & \text{s.t. } \mathbb{P} \left(\mathbf{A}_i \mathbf{x} + \tilde{\mathbf{A}}_i \mathbf{u}(\omega) \leq \mathbf{b}_i, \mathbf{f}_i(\mathbf{x}, \mathbf{u}(\omega)) \leq \mathbf{d}_i \right) \geq \Delta_i \\ & \forall i \in \{0, 1, \dots, N_C\} \end{aligned}$$

where:

- $\mathbf{x} \in \mathbb{R}^{N_R} \times \mathbb{Z}^{N_Z}$ is vector of decision variables;
- $c : \mathbb{R}^{N_R} \times \mathbb{Z}^{N_Z} \times \mathbb{R}^{N_U} \rightarrow \mathbb{R}$ is an objective function;
- \mathbf{A}_i are $N_{i,lin} \times (N_R + N_Z)$ matrices;
- $\tilde{\mathbf{A}}_i$ are $N_{i,lin} \times N_U$ matrices;
- $\mathbf{f}_i : \mathbb{R}^{N_R} \times \mathbb{Z}^{N_Z} \times \mathbb{R}^{N_U} \rightarrow \mathbb{R}^{N_{i,nlin}}$ are vector-valued measurable functions mapping from the space of the decision variables and random variables; and
- $\mathbf{b}_i \in \mathbb{R}^{N_{i,lin}}$ and $\mathbf{d}_i \in \mathbb{R}^{N_{i,nlin}}$ are vectors of constants.

The problem description assumes an underlying probability distribution for the random variables, although only samples are available. The problem requires assignments to decision variables which optimises expected utility while bounding the probability of constraint violation.

2.1 Initial Approach

In the initial approach, I will address the problem of uncertainty representation by constructing uncertainty sets from data rather than fitting distributions. Unlike prior work in robust programming, I will derive multiple uncertainty sets to be reasoned over during optimisation.

One way of providing probabilistic guarantees is to construct robust sets which cover a set of outcome with a sufficient probability mass, and then performing robust optimisation. The key idea is thus to estimate the location of the lower and upper quantiles. Due to the finite set of samples, the methods must be robust with respect to variations in sampling. I will thus also provide a *confidence* for estimates, a bound on the probability of estimates being incorrect.

Rather than directly fitting a cumulative distribution function from the data, I extend results from confidence intervals

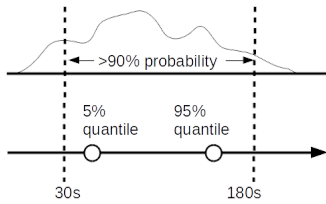


Figure 1: Probabilistic guarantee from bounds on quantiles.

for quantiles using order-statistics [Gibbons and Chakraborti, 2011] to find lower- and upper-bounds for the values of the quantiles, with the required confidence. For ordered sample u_k out of N samples, with probability α , it is:

- greater than the $\epsilon = 1 - B^{-1}(\alpha, N - k + 1, k)$ lower quantile b_ϵ^l ;
- smaller than the $\epsilon = 1 - B^{-1}(\alpha, k, N - k + 1)$ upper quantile b_ϵ^u ;

for $B^{-1}(\alpha, a, b)$ the α -quantile of the $Beta(a, b)$ distribution.

Fixing confidence, we can thus construct a finite number of bounds for quantile values, drawn from sampled values. The bounds on quantiles can be used in the same way as discrete cumulative distribution functions to provide probabilistic guarantees. An example is given in Figure 1, in which with 99% confidence the 5% quantile traversal is greater than 30 seconds and the 95% quantile is less than 180 seconds. Then, with 99% confidence the traversal takes between 30 seconds to 180 seconds with probability at least 90%.

I have implemented the algorithm *Fantana* which constructs the lower- and upper-bounds from marginal samples. I have tested the correctness of the method empirically on a data-set of 48361 samples of Boston subway traversal times between two adjacent stops, Davis Square and Porter Square. For each trial, I wished to find a lower-bound on the lower $\epsilon = 5\%$ quantile, with confidence $\alpha = 1\%$. Intuitive, I wanted to find a traversal time, so that only 5% of traversals will be shorter, with confidence 99%.

For each trial, I randomly assigned 4000 of the data points to for the training data set, and the rest to the test set. I then ordered the samples in the training set, and found sample points u_k , where k was the largest k such that $1 - B^{-1}(\alpha, k, N - k + 1) \leq 0.05$, for $N = 4000$ the size of the training set. For comparison, I also fit a distribution using the kernel density estimation implementation from the Scipy library [Jones et al., 2001]. I then noted the cumulative density at u_k according to the fitted distribution. To test the correctness of the bound, I then determined the proportion of data points in the test set which was smaller than the chosen bound.

Averaging over 10000 trials, the average proportion of test data points smaller than the chosen lower-bound u_k 's was 3.4%, suggesting that the method is not overly conservative. Further, the proportion of test data points smaller than the lower bound was less than 5% in all 10000 trials, showing correctness with respect to the 99% confidence bound.

However, in 989 trials, the proportion of test data points was greater than the cumulative density at u_k fitted using kernel density estimation. Thus, kernel density estimation provided an incorrect bound in 9.89% of the trials.

The contributions to CC-CP are as follows: 1) Explicit confidence on the estimates enforces robustness to sampling variations; 2) constructing multiple quantile bounds allows agents to choose robust sets with greater flexibility than existing robust optimisation methods; and 3) discrete choices for robust set bounds allows agents to leverage techniques in discrete search instead of optimising over nonlinear density functions.

3 Future Work

Fantana currently constructs multiple uncertainty sets with correct probabilistic guarantees. In subsequent work, I will demonstrate faster computation for chance-constrained constraint programming when the constructed sets are used for uncertainty representation. *Fantana* will be used to construct robust sets for scheduling under probabilistic uncertainty with chance-constrained probabilistic simple temporal problems (cc-pSTPs) [Fang et al., 2014], using data for traversal times and wait times at station a subway line.

The key to scalable solutions to cc-pSTPs is the decomposition of the problem into the risk management and the temporal consistency checking subproblems. The risk management master constructs uncertainty sets for durations to generate a candidate deterministic temporal subproblem. The subproblem is checked with use efficient scheduling algorithms, and inconsistency are explained by returning conflicts to the master in the form of linear constraints. I will prove that these linear constraints corresponds to Bender's cuts for feasibility, and derive corresponding optimality cuts for cc-pSTPs. I will then generalise cut learning for chance-constrained constraint programs of the form in Problem 1. Cut learning will also reduce over-conservatism by identifying restrictions over the outcomes of uncontrollable variables, as bounds on functions of uncontrollable variable outcomes. Measurable functions can also be considered random variables, and uncertainty sets can be iteratively constructed with *Fantana* for subsequent risk allocation.

References

[Bertsimas et al., 2017] Dimitris Bertsimas, Vishal Gupta, and Nathan Kallus. Data-driven robust optimization. *Mathematical Programming*, pages 1–58, 2017.

[Fang et al., 2014] Cheng Fang, Peng Yu, and Brian C Williams. Chance-constrained probabilistic simple temporal problems. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 2264–2270. AAAI Press, 2014.

[Gibbons and Chakraborti, 2011] Jean Dickinson Gibbons and Subhabrata Chakraborti. *Nonparametric statistical inference*. Springer, 2011.

[Gill et al., 2005] Philip E Gill, Walter Murray, and Michael A Saunders. Snpot: An sqp algorithm for large-scale constrained optimization. *SIAM review*, 47(1):99–131, 2005.

[Jones et al., 2001] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001. [Online; accessed 2017-03-06].

[Wächter and Biegler, 2006] Andreas Wächter and Lorenz T Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106(1):25–57, 2006.