On distributionally robust joint chance-constrained problems

Weijun Xie *

Advisor: Shabbir Ahmed

Introduction: A chance constrained optimization problem involves constraints with stochastic data that are required to be satisfied with a pre-specified probability. When the underlying distribution of the stochastic data is not known precisely, an often used model is to require the chance constraints to hold for all distributions in a given family. Such a problem is known as a distributionally robust chance constrained problem (DRCCP). We consider mixed integer linear DRCCP problems of the form:

$$
\begin{aligned}
v^* = \min\ & c^\top x, && \text{(1a)}\\
\text{s.t.}\ & x \in S, && \text{(1b)}\\
& \mathbb{P}[A(x)\xi \le b(x)] \ge 1-\epsilon \ \ \forall\, \mathbb{P} \in \mathcal{P}. && \text{(1c)}
\end{aligned}
$$

In the above formulation we seek a decision vector x to minimize a linear objective $c^\top x$ subject to a set of deterministic mixed integer constraints defined by $S = \{x \in \mathbb{R}^{n-\tau} \times \mathbb{Z}^\tau : Dx \ge d\}$, and distributionally robust chance constraints (1c) that are required to hold with a probability of $1-\epsilon$ where $\epsilon \in (0,1)$ is a risk parameter. The stochastic data in (1c) is modeled by a random vector ξ supported on a nonempty compact set $\Xi \subset \mathbb{R}^m$, and the ambiguity set $\mathcal{P}$ defines a family of probability distributions $\mathbb{P}$ on Ξ. The constraint system in (1c) is defined by the technology matrix $A(x) = (a_1(x), \ldots, a_I(x))^\top$ and the right-hand side vector $b(x) = (b_1(x), \ldots, b_I(x))^\top$ which are affine in x, i.e. , $a_i(x) = A^i x + a^i$ and $b_i(x) = B^i x + b^i$ with $A^i = \mathbb{R}^{m\times n}, a^i \in \mathbb{R}^m, B^i \in \mathbb{R}^n, b^i \in \mathbb{R}$ for each $i \in [I]$. Note that when $I = 1$, (1c) is called "*single*" chance constraint while it is a "*joint*" chance constraint if $I > 1$.

Existing Work: Even when $\mathcal{P}$ is a singleton and $|I| = 1$, i.e. a single unambiguous chance constraint system, the feasible region of (1c) is highly nonconvex. An important line of work has been to develop convex inner approximation of this feasible region to obtain safe tractable approximations of chance constrained problems. A noteworthy work is Nemirovski and Shapiro (2006) where it is proved that the so-called conditional value-at-risk (**CVaR**) constraint provides the tightest convex approximation to an unambiguous single chance constraint. For joint chance constraints, Nemirovski and Shapiro (2006) suggested using Bonferroni's inequality to decompose (1c) into $|I|$ single chance constraints whose sum of risk parameters is no larger than ϵ, then for each single chance constraint, the **CVaR** approximation could be directly applied. However, such an approach has been shown to be weak (Chen et al., 2010; Zymler et al., 2013). Chen et al. (2010) improved the approach by scaling each uncertain constraint in a joint chance constraint system with a positive number and aggregating them into a single constraint, and then they were able to provide a conservative approximation of the joint chance constraint system via second order cone program (SOCP). Zymler et al. (2013) combined the **CVaR** approximation approach of Nemirovski and Shapiro (2006) with the scaling approach of Chen et al. (2010) to develop nearly tight convex approximations of DRCCP when $\mathcal{P}$ is given by first and second moment constraints.

Research Objective: We propose to generalize the construction of Zymler et al. (2013) for building convex approximations of DRCCP under general ambiguity sets $\mathcal{P}$ and explore the tightness of such approximations. In particular we consider arbitrary convex moment ambiguity set defined as

$$
\mathcal{P}^M = \left\{\mathbb{P} \in \mathcal{P}_0\left(\Xi\right) : \mathbb{E}_{\mathbb{P}}[\phi_t(\xi)] \ge g_t, t = 1, \ldots, T\right\}, \tag{2}
$$

where $\mathcal{P}_0\left(\Xi\right)$ denotes the set of all probability measures on Ξ with a sigma algebra $\mathcal{F}$, $\phi_t(\xi)$ is real valued convex function on $(\Xi, \mathcal{F})$ for each $t \in [T]$. This generalization allows much richer distributional information than that specified by first and second moments. For example, it allows consideration of moment information on indirect observations or convolutions of the uncertain parameters.

*PhD student. Email: wxie33@gatech.edu.

Approach and Preliminary Results: We proceed as in Chen et al. (2010) and Zymler et al. (2013) by first scaling each constraint $i \in [I]$ in the joint chance constraint system by a positive scalar α_i, aggregating them to single chance constraint, and building a $\mathbf{CVaR}$ approximation of (1) which takes the form

$$Z_{\mathbf{CVaR}} := \left\{ x \in S : \alpha \in \mathbb{R}^I_{++}, \inf_{\beta \in \mathbb{R}} \left\{ \beta + \frac{1}{\epsilon} \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[\left(\max_{i \in [I]} \left\{ \alpha_i \left(a_i(x)^\top \xi - b_i(x) \right) \right\} - \beta \right)_+ \right] \right\} \leq 0 \right\}. \tag{3}$$

Let Z denote the feasible region of (1), then $Z_{\mathbf{CVaR}} \subseteq Z$. Under appropriate regularity conditions to ensure of strong duality of semi-infinite programs, we can show that $Z_{\mathbf{CVaR}}$ is almost equal to Z; i.e.,

Theorem 1. *Suppose that $\mathcal{P} = \mathcal{P}^M$ then (under some regularity conditions)*

$$Z_{int} \subseteq Z_{\mathbf{CVaR}} \subseteq Z,$$

where $Z_{int} := \{x \in S : \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[A(x)\xi < b(x)] \geq 1 - \epsilon\}$.

The key to this result is to show that the scaling approach of Chen et al. (2010) is in some sense optimal, and then to exploit the duality structure of semi-infinite linear optimization over $\mathcal{P}^M$.

By dualizing the inner maximization problem, we can reformulate $Z_{\mathbf{CVaR}}$ as a deterministic counterpart involving constraints that are convex in x. Thus we have a nearly tight deterministic reformulation of DRCCP. However, this reformulation has bilinear terms (i.e., $\alpha_i a_i(x), \alpha_i b_i(x)$ for each $i \in I$ in (3)), which are usually nonconvex jointly in α and x. We will investigate convexification approaches for these bilinear terms. Some possible directions are outlined below.

- When x are binary variables (i.e., $S \subseteq \{0,1\}^n$), then we could introduce new variables to represent these bilinear terms and linearize them via McCormick inequalities. Such a reformulation technique will yield a convex mixed-integer program, which is an exact reformulation of $Z_{\mathbf{CVaR}}$. The precise construction and tightness needs further study.
- When there are is one moment constraint in the ambiguity set $\mathcal{P}^M$ (i.e., $T = 1$), then it appears that $Z_{\mathbf{CVaR}}$ could be reformulated as a convex set by projecting out the scaling variables α, hence, the bilinear terms will vanish. This will provide a tractable convex approximation.
- If the stochastic data ξ does not interact with decision variables x (i.e., in (1c), $a_i(x)$ is constant for each $i \in [I]$), and if $\phi_t(\xi)$ is a positively homogeneous function for each $t \in [T]$ while Ξ is a closed convex cone, then we conjecture that $Z_{\mathbf{CVaR}}$ is equivalent to a convex hull of two disjunctive convex sets. This implies that in order to solve (1) under moment ambiguity set, we can solve two convex optimization problems with respect to two convex sets separately, and then choose the smaller optimal objective value.

We also propose to study DRCCP with joint constraint under data driven ambiguity set constructing from empirical distribution; for instance, ambiguity set can be built on all of probability distributions whose distance with some known empirical distribution is within a threshold. When the empirical distribution is discrete, we will investigate Wasserstein metric based ambiguity set, while when empirical distribution is continuous, the ϕ- divergence metric is more appropriate. A similar dualization technique from moment ambiguity set can be applied when reformulating (1c) into its deterministic counterpart. Under Wasserstein metric ambiguity set, we believe that Z can also be tightly approximated by another deterministic reformulation with bilinear terms for which convexification approaches need to be developed. For ϕ- divergence ambiguity set, we can show that (1c) is equivalent to a regular chance constrained problem with empirical distribution as its underlying probability distribution and possibly a smaller risk parameter. Thus the standard $\mathbf{CVaR}$ approximation suffices.

References

Chen, W., Sim, M., Sun, J., and Teo, C.-P. (2010). From cvar to uncertainty set: Implications in joint chance-constrained optimization. *Operations research*, 58(2):470–485.

Nemirovski, A. and Shapiro, A. (2006). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996.

Zymler, S., Kuhn, D., and Rustem, B. (2013). Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198.