A stochastic approximation method for chance-constrained nonlinear programs

Rohit Kannan

UW-Madison

December 10, 2018

Joint work with Jim Luedtke

Outline

Formulation

$$\nu^* := \min_{x \in X} f(x) \qquad \text{(CCP)}$$
$$\text{s.t. } \mathbb{P}\{g(x,\xi) \leq 0\} \geq 1 - \alpha.$$

- Assume:
  - $X \subset \mathbb{R}^n$ is nonempty, compact, and convex
  - $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and quasiconvex
  - $g : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is continuously differentiable
  - $\xi \sim \mathcal{P}$ is continuous with support $\Xi \subset \mathbb{R}^d$
  - Other relatively mild technical assumptions ...

Formulation

$$\nu^* := \min_{x \in X} \; f(x) \qquad \text{(CCP)}$$
$$\text{s.t. } \mathbb{P}\left\{g(x,\xi) \leq 0\right\} \geq 1 - \alpha.$$

- Assume:
  - $X \subset \mathbb{R}^n$ is nonempty, compact, and convex
  - $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and quasiconvex
  - $g : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is continuously differentiable
  - $\xi \sim \mathcal{P}$ is continuous with support $\Xi \subset \mathbb{R}^d$
  - Other relatively mild technical assumptions ...
- Can model joint chance constraints, deterministic nonconvex constraints, and some recourse structures

Central hypothesis

Main idea #1

In many cases, decision makers are interested in generating the efficient frontier of optimal objective function value (ν^*) versus risk level (α) rather than simply solving (CCP) for a single prespecified risk level so that they can make a more informed decision.

— Rengarajan and Morton [2009], Luedtke [2014]

Central hypothesis

Main idea #1

> In many cases, decision makers are interested in generating the efficient frontier of optimal objective function value (ν^*) versus risk level (α) rather than simply solving (CCP) for a single prespecified risk level so that they can make a more informed decision.

— Rengarajan and Morton [2009], Luedtke [2014]

- Efficient frontier of (CCP) can be recovered by solving the following stochastic optimization problem:

$$\min_{x \in X} \mathbb{P}\left\{g(x,\xi) \not\leq 0\right\} \quad \equiv \quad \min_{x \in X} \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \qquad \text{(SP)}$$

$$\text{s.t. } f(x) \leq \nu. \qquad\qquad \text{s.t. } f(x) \leq \nu.$$

  - $\mathbb{1}\left[\cdot\right]$ denotes the l.s.c. step function

Main idea #2

- Approximate the efficient frontier of (CCP) using:

$$\min_{x \in X} \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \quad \approx \quad \min_{x \in X} \mathbb{E}\left[\max\left[\phi_k\left(g(x,\xi)\right)\right]\right] \qquad (\mathrm{APP}_k)$$

$$\text{s.t. } f(x) \leq \nu. \qquad\qquad\qquad\qquad \text{s.t. } f(x) \leq \nu.$$

$\{\phi_k\}$ is a 'convergent' seq. of smooth approx. of the step function.

Main idea #2

- Approximate the efficient frontier of (CCP) using:

$$\min_{x \in X} \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \quad \approx \quad \min_{x \in X} \mathbb{E}\left[\max\left[\phi_k\left(g(x,\xi)\right)\right]\right] \qquad (\mathrm{APP}_k)$$

$$\text{s.t. } f(x) \le \nu. \qquad\qquad \text{s.t. } f(x) \le \nu.$$

$\{\phi_k\}$ is a 'convergent' seq. of smooth approx. of the step function.

- (APP$_k$) can be solved using stochastic gradient methods.

Main idea #2

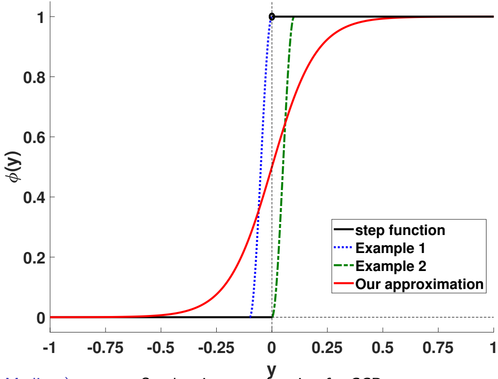
- Approximate the efficient frontier of (CCP) using:

$$\min_{x \in X} \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \quad \approx \quad \min_{x \in X} \mathbb{E}\left[\max\left[\phi_k\left(g(x,\xi)\right)\right]\right] \quad (\mathrm{APP}_k)$$

$$\text{s.t. } f(x) \leq \nu. \qquad\qquad \text{s.t. } f(x) \leq \nu.$$

$\{\phi_k\}$ is a 'convergent' seq. of smooth approx. of the step function.

- (APP_k) can be solved using stochastic gradient methods.



Outline

Scenario approximation

- Sample N random realizations of ξ from $\mathcal{P}$
- Solve the following scenario problem to local optimality using a cutting-plane approach:

$$\hat{x}_N \in \operatorname*{arg\,min}_{x \in X} \; f(x)$$
$$\text{s.t.} \quad g(x, \xi_j) \leq 0, \quad \forall j \in \{1, \cdots, N\}.$$

- Estimate the risk level $\hat{\alpha}_N := \mathbb{P}\{g(\hat{x}_N, \xi) \nleq 0\}$ using an independent Monte Carlo sample to determine the point $(\hat{\alpha}_N, f(\hat{x}_N))$ on the approximation of the efficient frontier

- References: Calafiore and Campi [2005], Campi and Garatti [2011]

Existing smoothing-based approaches

- Approximate the solution of (CCP) using:

$$\begin{aligned} &\min_{x \in X} \; f(x) && \approx \quad \min_{x \in X} \; f(x) \\ &\text{s.t. } \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \le \alpha. && \quad \text{s.t. } \mathbb{E}\left[\max\left[\phi_k\left(g(x,\xi)\right)\right]\right] \le \alpha. \end{aligned}$$

where $\{\phi_k\}$ is a 'convergent' sequence of conservative smooth approximations of the step function.

- Solution of each approximation yields a feasible solution to the original chance-constrained program

Existing smoothing-based approaches

- Approximate the solution of (CCP) using:

$$\min_{x \in X} \ f(x) \quad \text{s.t. } \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \leq \alpha. \qquad \approx \qquad \min_{x \in X} \ f(x) \quad \text{s.t. } \mathbb{E}\left[\max\left[\phi_k\left(g(x,\xi)\right)\right]\right] \leq \alpha.$$

where $\{\phi_k\}$ is a 'convergent' sequence of conservative smooth approximations of the step function.

- Solution of each approximation yields a feasible solution to the original chance-constrained program
- Currently, the only known approach for solving optimization problems with nonconvex expectation constraints is via sample average approximation

- References: Hong et al. [2011], Shan et al. [2014, 2016], Geletu et al. [2017], Cao and Zavala [2017], Adam et al. [2018]

Pros and cons of exterior sampling approaches

Pros

- Can use off-the-shelf solvers
- Possess strong theoretical guarantees

Cons

- May find spurious local optima
- May need to solve large-scale NLPs

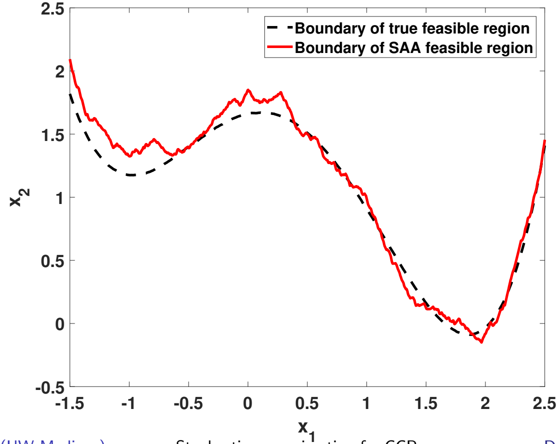
- We propose the first interior sampling approach for solving chance-constrained problems

Sampling may lead to spurious local minima

- Consider the following chance constraint [Curtis et al., 2018]:

$$\mathbb{P}\left\{0.25x_1^4 - \frac{1}{3}x_1^3 - x_1^2 + 0.2x_1 - 19.5 + \xi_1 x_1 + \xi_1\xi_0 \leq x_2\right\} \geq 0.95,$$

where $\xi_1 \sim U(-3,3)$ and $\xi_0 \sim U(-12,12)$ are independent.



Outline

Main theoretical results

- Let $\{\phi_k\}$ be a 'convergent' sequence of smooth approximations of the step function.

$$\min_{x \in X} \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \quad \text{(SP)} \qquad \min_{x \in X} \mathbb{E}\left[\max\left[\phi_k\left(g(x,\xi)\right)\right]\right] \quad (\text{APP}_k)$$

$$\text{s.t. } f(x) \le \nu. \qquad\qquad \text{s.t. } f(x) \le \nu.$$

Theorem

Every limit point of a sequence of global solutions to the approximations (APP_k) *is a global solution to Problem* (SP)

Tentative Theorem

Every limit point of a sequence of stationary solutions to the approximations (APP_k) is a stationary solution to Problem (SP)

Proposal for approximating the efficient frontier

- Input: initial guess $\hat{x}^0 \in X$, sequence of objective bounds $\{\bar{\nu}^k\}$ (determined by solving a scenario approximation problem), and lower bound on risk level $\alpha_{low} \in (0, 1)$
- Output: pairs $\{(\bar{\nu}^i, \bar{\alpha}^i)\}$ of objective values and risk levels used to approximate the efficient frontier ($+$ solutions $\{\bar{x}^i\}$)

Proposal for approximating the efficient frontier

- Input: initial guess $\hat{x}^0 \in X$, sequence of objective bounds $\{\bar{\nu}^k\}$ (determined by solving a scenario approximation problem), and lower bound on risk level $\alpha_{low} \in (0, 1)$
- Output: pairs $\{(\bar{\nu}^i, \bar{\alpha}^i)\}$ of objective values and risk levels used to approximate the efficient frontier ($+$ solutions $\{\bar{x}^i\}$)
- Solve: sequence of problems (APP_k) with objective bound $\bar{\nu}^k$ using the projected stochastic subgradient method [Davis and Drusvyatskiy, 2018]
- Estimate: risk level $\bar{\alpha}^k$ of the best found solution using an independent Monte Carlo sample to determine the point $(\bar{\nu}^k, \bar{\alpha}^k)$ on our approximation of the efficient frontier
- Terminate: when $\bar{\alpha}^k \leq \alpha_{low}$

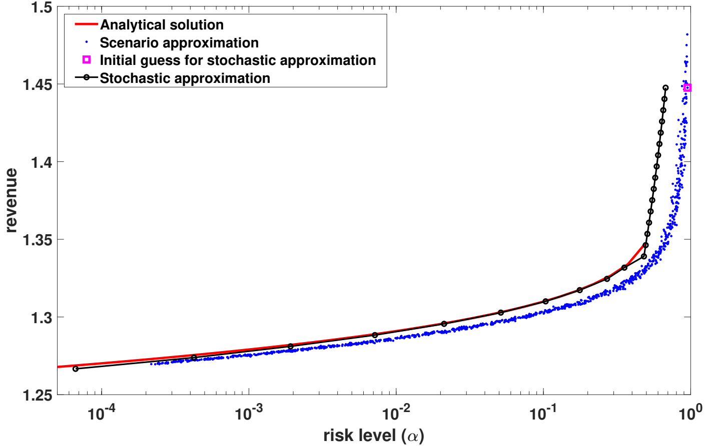
Outline

Portfolio optimization [Ben-Tal et al., 2009]

$$\max_{t,\, x \in X} \; t$$

$$\text{s.t. } \mathbb{P}\left\{\xi^{\mathsf{T}} x \geq t\right\} \geq 1 - \alpha,$$

where $X := \left\{x \in \mathbb{R}^{1000}_{+} : \sum_i x_i = 1\right\}$, ξ_i are independent Normal.



Resource planning [Luedtke, 2014]

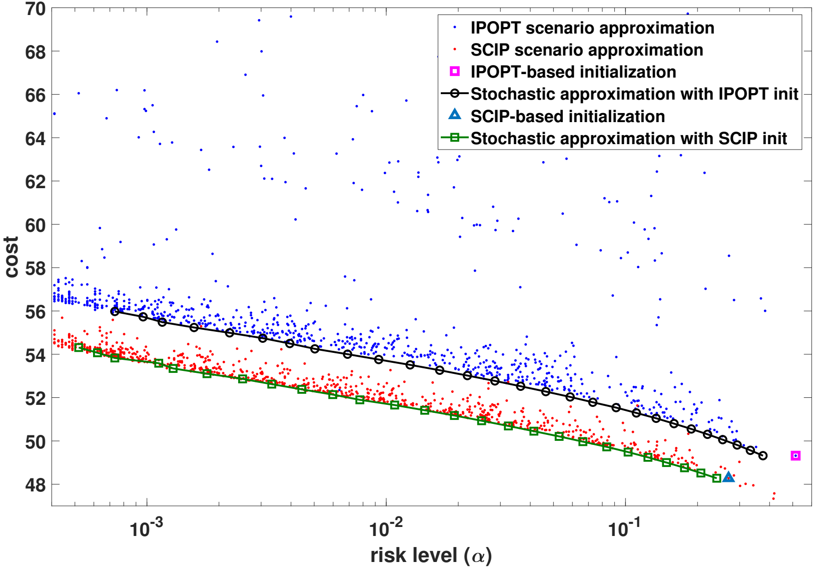
$$\min_{x \in \mathbb{R}^{20}_+} \quad c^\mathsf{T} x$$
$$\text{s.t.} \quad \mathbb{P}\left\{x \in R(\lambda, \rho)\right\} \geq 1 - \alpha,$$

where

$$R(\lambda,\rho) = \left\{ x \in \mathbb{R}^{20}_+ \; : \; \exists y \in \mathbb{R}^{20\times 30}_+ \text{ s.t. } \sum_{j=1}^{30} y_{ij} \leq \rho_i x_i^2, \; \forall i \in \{1, \cdots, 20\}, \right.$$
$$\left. \sum_{i=1}^{20} \mu_{ij} y_{ij} \geq \lambda_j, \; \forall j \in \{1, \cdots, 30\} \right\}.$$

- x_i: quantity of resource i, c_i: unit cost of resource i
- y_{ij}: amount of resource i allocated to customer type j
- $\rho_i \in (0, 1]$: random yield of resource i
- $\lambda_j \geq 0$: random demand of customer type j
- $\mu_{ij} \geq 0$: service rate of resource i for customer type j

Resource planning



Outline

Summary and Open questions

- Proposed a stochastic approximation approach for generating the efficient frontier of chance-constrained NLPs
- Computational results indicate that our proposal consistently determines better approximations of the efficient frontier than existing approaches in reasonable computation times

- Open questions
  - Extension to multiple sets of joint chance constraints
  - Handle deterministic nonconvex constraints more naturally
  - Theory for randomized constraint projection techniques to reduce effort spent on projections
  - Extension to distributionally robust chance constraints

References I

Lukáš Adam, Martin Branda, Holger Heitsch, and René Henrion. Solving joint chance constrained problems using regularization and Benders decomposition. *Annals of Operations Research*, pages 1–27, 2018. doi: 10.1007/s10479-018-3091-9.

Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton University Press, 2009.

Giuseppe Calafiore and Marco C Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, 2005.

Marco C Campi and Simone Garatti. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *Journal of Optimization Theory and Applications*, 148(2):257–280, 2011.

Yankai Cao and Victor Zavala. A sigmoidal approximation for chance-constrained nonlinear programs. `http://www.optimization-online.org/DB_FILE/2017/10/6236.pdf`, 2017.

Frank E Curtis, Andreas Wachter, and Victor M Zavala. A sequential algorithm for solving nonlinear optimization problems with chance constraints. *SIAM Journal on Optimization*, 28(1):930–958, 2018.

Damek Davis and Dmitriy Drusvyatskiy. Stochastic subgradient method converges at the rate $O\left(k^{-1/4}\right)$ on weakly convex functions. *arXiv preprint arXiv:1802.02988*, 2018.

References II

Abebe Geletu, Armin Hoffmann, Michael Kloppel, and Pu Li. An inner-outer approximation approach to chance constrained optimization. *SIAM Journal on Optimization*, 27(3):1834–1857, 2017.

L Jeff Hong, Yi Yang, and Liwei Zhang. Sequential convex approximations to joint chance constrained programs: A Monte Carlo approach. *Operations Research*, 59 (3):617–630, 2011.

James Luedtke. A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support. *Mathematical Programming*, 146(1-2):219–244, 2014.

Tara Rengarajan and David P Morton. Estimating the efficient frontier of a probabilistic bicriteria model. In *Winter Simulation Conference*, pages 494–504, 2009.

F Shan, XT Xiao, and LW Zhang. Convergence analysis on a smoothing approach to joint chance constrained programs. *Optimization*, 65(12):2171–2193, 2016.

Feng Shan, Liwei Zhang, and Xiantao Xiao. A smoothing function approach to joint chance-constrained programs. *Journal of Optimization Theory and Applications*, 163(1):181–199, 2014.