A Stochastic Approximation Method for Approximating the Efficient Frontier of Chance-Constrained Nonlinear Programs

Mathematical Programming Computation, 13(4), pp. 705-751

Rohit Kannan

Center for Nonlinear Studies Postdoctoral Fellow
Los Alamos National Laboratory

November 8, 2021

Joint work with Jim Luedtke (UW-Madison)

Funding: DOE (MACSER Project), Center for Nonlinear Studies

Chance-Constrained Programming

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

- Introduced by Charnes et al. (1958); Charnes and Cooper (1959)
- Extended by Miller and Wagner (1965) to include joint chance constraints and by Prékopa (1970) to the nonlinear setting

Chance-Constrained Programming

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

- Introduced by Charnes et al. (1958); Charnes and Cooper (1959)
- Extended by Miller and Wagner (1965) to include joint chance constraints and by Prékopa (1970) to the nonlinear setting
- Applications:
  - Process optimization (Li et al., 2008)
  - Infrastructure networks (Gotzes et al., 2016; Roald et al., 2013)
  - Portfolio optimization (Shapiro et al., 2009)

Chance-Constrained Programming

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

- Introduced by Charnes et al. (1958); Charnes and Cooper (1959)
- Extended by Miller and Wagner (1965) to include joint chance constraints and by Prékopa (1970) to the nonlinear setting
- Applications:
  - Process optimization (Li et al., 2008)
  - Infrastructure networks (Gotzes et al., 2016; Roald et al., 2013)
  - Portfolio optimization (Shapiro et al., 2009)
- Challenges:
  1. Feasible region can be nonconvex even if g is affine in (x, ξ)
  2. Checking feasibility involves multi-dimensional integration

Formulation

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

- 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
  - Other relatively mild technical assumptions . . .

Formulation

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

- 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
  - Other relatively mild technical assumptions . . .
- Do not assume:
  - Distribution of the random vector ξ (only need i.i.d. samples)
  - Structure of the vector-valued function g

Formulation

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

- 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
  - Other relatively mild technical assumptions ...
- Do not assume:
  - Distribution of the random vector ξ (only need i.i.d. samples)
  - Structure of the vector-valued function g
- Can model joint chance constraints, deterministic nonconvex constraints, and some models with recourse

Solution Approaches

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

- Deterministic equivalent (Prékopa, 1995; Lagoa et al., 2005)
  - Reformulate as a convex program
  - Strong assumptions on form of g, distribution of ξ

Solution Approaches

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

- Deterministic equivalent (Prékopa, 1995; Lagoa et al., 2005)
  - Reformulate as a convex program
  - Strong assumptions on form of g, distribution of ξ
- Convex inner-approximations (Rockafellar and Uryasev, 2000; Nemirovski and Shapiro, 2006)
  - Inner-approximate the feasible region with a convex set
  - Can be solved efficiently, but often yield conservative solutions

Solution Approaches

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

- Deterministic equivalent (Prékopa, 1995; Lagoa et al., 2005)
  - Reformulate as a convex program
  - Strong assumptions on form of g, distribution of ξ
- Convex inner-approximations (Rockafellar and Uryasev, 2000; Nemirovski and Shapiro, 2006)
  - Inner-approximate the feasible region with a convex set
  - Can be solved efficiently, but often yield conservative solutions
- Several other approaches, e.g., Norkin (1993); Andrieu et al. (2007); Lepp (2009); van Ackooij and Henrion (2014, 2017); Curtis et al. (2018); Peña-Ordieres et al. (2020)

Scenario Approximation

- Draw a fixed sample $\{\xi^i\}_{i=1}^N$ of the random vector
- Solve the scenario approximation problem (Calafiore and Campi, 2005; Campi and Garatti, 2011)

$$\hat{x}_N \in \underset{x \in X}{\arg\min} \ f(x)$$
$$\text{s.t.} \ \ g(x, \xi^i) \leq 0, \quad \forall i \in \{1, \ldots, N\}$$

Scenario Approximation

- Draw a fixed sample $\{\xi^i\}_{i=1}^N$ of the random vector
- Solve the scenario approximation problem (Calafiore and Campi, 2005; Campi and Garatti, 2011)

$$\hat{x}_N \in \underset{x \in X}{\arg\min} \; f(x)$$
$$\text{s.t.} \;\; g(x, \xi^i) \leq 0, \quad \forall i \in \{1, \dots, N\}$$

- Estimate $\mathbb{P}\{g(\hat{x}_N, \xi) \leq 0\}$ and tune sample size N accordingly
- Theory on sample size N required to ensure that $\hat{x}_N$ is feasible to (CCP) with high probability

Sample Average Approximation

- Draw a fixed sample $\{\xi^i\}_{i=1}^N$ of the random vector
- Solve the SAA problem (Luedtke and Ahmed, 2008)

$$\hat{x}_N \in \underset{x \in X}{\arg\min} \; f(x)$$

$$\text{s.t.} \quad \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\left[g(x, \xi^i)\right] \leq \gamma$$

  for some $\gamma \in [0, \alpha)$, where $\mathbb{1}[\cdot]$ is the l.s.c. step function
- Can reformulate as an MI(N)LP, but typically require tailored approaches for efficient solution

Sample Average Approximation

- Draw a fixed sample $\{\xi^i\}_{i=1}^N$ of the random vector
- Solve the SAA problem (Luedtke and Ahmed, 2008)

$$\hat{x}_N \in \operatorname*{arg\,min}_{x \in X} \; f(x)$$

$$\text{s.t.} \quad \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\left[g(x, \xi^i)\right] \leq \gamma$$

  for some $\gamma \in [0, \alpha)$, where $\mathbb{1}[\cdot]$ is the l.s.c. step function
- Can reformulate as an MI(N)LP, but typically require tailored approaches for efficient solution
- Estimate $\mathbb{P}\{g(\hat{x}_N, \xi) \leq 0\}$ and tune γ, N accordingly
- Theory on sample size N required to ensure that $\hat{x}_N$ is feasible to (CCP) with high probability

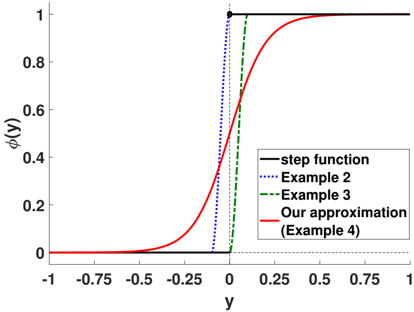
Smooth Approximation of SAA

- Recent renewed interest in smooth approximations (Hong et al., 2011; Geletu et al., 2017; Cao and Zavala, 2017)

$$\hat{x}_N \in \underset{x \in X}{\arg\min} \; f(x)$$

$$\text{s.t.} \quad \frac{1}{N}\sum_{i=1}^{N} \phi(g(x,\xi^i)) \leq \gamma$$

where $\phi(\cdot)$ is a smooth approximation of $\mathbb{1}[\cdot]$

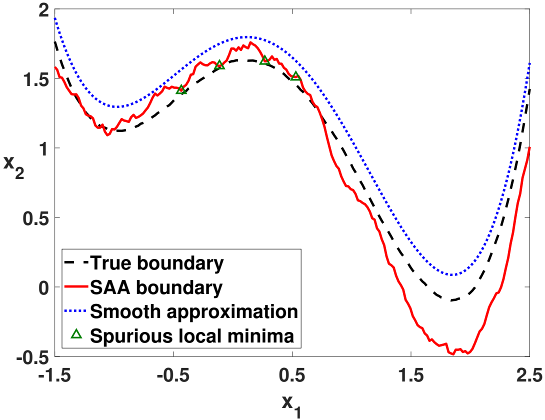


Exterior Sampling may lead to Spurious Local Minima

- Consider the following chance constraint (Curtis et al., 2018):

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

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



The Efficient Frontier of Risk versus Reward

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

Decision makers are often interested in generating the efficient frontier of optimal objective value (ν^*_α) versus risk level (α) rather than simply solving (CCP) for a single prespecified risk level.

— Rengarajan and Morton (2009); Luedtke (2014)

The Efficient Frontier of Risk versus Reward

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

Decision makers are often interested in generating the efficient frontier of optimal objective value (ν_α^*) versus risk level (α) rather than simply solving (CCP) for a single prespecified risk level.

— Rengarajan and Morton (2009); Luedtke (2014)

- Efficient frontier can be recovered by solving

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

Approximating the Efficient Frontier efficiently

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

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

where $\phi_k(\cdot) \to \mathbb{1}\left[\cdot\right]$ is a sequence of smooth approximations

Approximating the Efficient Frontier efficiently

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

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

where $\phi_k(\cdot) \to \mathbb{1}\left[\cdot\right]$ is a sequence of smooth approximations

- (APP_ν) can be solved using stochastic subgradient methods

Approximating the Efficient Frontier efficiently

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

where $\phi_k(\cdot) \to \mathbb{1}\left[\cdot\right]$ is a sequence of smooth approximations

---

Projected Stochastic Subgradient (Davis and Drusvyatskiy, 2018)

Input: Initial guess $x_1 \in X$, number of iterations T, mini-batch size M, and step length $\gamma > 0$

Approximating the Efficient Frontier efficiently

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

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

where $\phi_k(\cdot) \to \mathbb{1}\left[\cdot\right]$ is a sequence of smooth approximations

---

Projected Stochastic Subgradient (Davis and Drusvyatskiy, 2018)

Input: Initial guess $x_1 \in X$, number of iterations T, mini-batch size M, and step length $\gamma > 0$

for $t = 1, \cdots, T$ **do**

end for

Approximating the Efficient Frontier efficiently

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

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

where $\phi_k(\cdot) \to \mathbb{1}\left[\cdot\right]$ is a sequence of smooth approximations

---

Projected Stochastic Subgradient (Davis and Drusvyatskiy, 2018)

Input: Initial guess $x_1 \in X$, number of iterations T, mini-batch size M, and step length $\gamma > 0$

for $t = 1, \cdots, T$ **do**

  for $j = 1, \cdots, M$ **do**

    Let ξ^j be a random observation of ξ

    Compute $g_{t,j} \in \partial_x \max\left[\phi_k\left(g(x_t, \xi^j)\right)\right]$

  end for

end for

Approximating the Efficient Frontier efficiently

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

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

where $\phi_k(\cdot) \to \mathbb{1}\left[\cdot\right]$ is a sequence of smooth approximations

---

Projected Stochastic Subgradient (Davis and Drusvyatskiy, 2018)

Input: Initial guess $x_1 \in X$, number of iterations T, mini-batch size M, and step length $\gamma > 0$

for $t = 1, \cdots, T$ **do**

  for $j = 1, \cdots, M$ **do**

    Let ξ^j be a random observation of ξ

    Compute $g_{t,j} \in \partial_x \max\left[\phi_k\left(g(x_t, \xi^j)\right)\right]$

  end for

  Let $x_{t+1} = \mathrm{Proj}_X\left(x_t - \gamma \frac{1}{M} \sum_{j=1}^{M} g_{t,j}\right)$

end for

Approximating the Efficient Frontier efficiently

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

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

where $\phi_k(\cdot) \to \mathbb{1}\left[\cdot\right]$ is a sequence of smooth approximations

---

Projected Stochastic Subgradient (Davis and Drusvyatskiy, 2018)

Input: Initial guess $x_1 \in X$, number of iterations T, mini-batch size M, and step length $\gamma > 0$

for $t = 1, \cdots, T$ **do**

  for $j = 1, \cdots, M$ **do**

    Let ξ^j be a random observation of ξ

    Compute $g_{t,j} \in \partial_x \max\left[\phi_k\left(g(x_t, \xi^j)\right)\right]$

  end for

  Let $x_{t+1} = \mathrm{Proj}_X\left(x_t - \gamma \frac{1}{M} \sum_{j=1}^{M} g_{t,j}\right)$

end for

Output: x_{T+1} or iterate with smallest estimated objective value

---

Outline of the Algorithm

Approximating the Efficient Frontier of (CCP)

Input: initial guess $x^0 \in X$, sequence of objective bounds $\{\nu^k\}$, and lower bound on risk level $\alpha_{low} \in (0, 1)$

Output: pairs $\{(\nu^i, \alpha^i)\}$ of objective values and risk levels used to approximate the efficient frontier, solutions $\{x^i\}$

Outline of the Algorithm

Approximating the Efficient Frontier of (CCP)

Input: initial guess $x^0 \in X$, sequence of objective bounds $\{\nu^k\}$, and lower bound on risk level $\alpha_{low} \in (0, 1)$

Output: pairs $\{(\nu^i, \alpha^i)\}$ of objective values and risk levels used to approximate the efficient frontier, solutions $\{x^i\}$

Preprocessing: determine suitably scaled sequence of smoothing functions $\{\phi_k\}$ and a sequence of step lengths

Outline of the Algorithm

Approximating the Efficient Frontier of (CCP)

Input: initial guess $x^0 \in X$, sequence of objective bounds $\{\nu^k\}$, and lower bound on risk level $\alpha_{low} \in (0, 1)$

Output: pairs $\{(\nu^i, \alpha^i)\}$ of objective values and risk levels used to approximate the efficient frontier, solutions $\{x^i\}$

Preprocessing: determine suitably scaled sequence of smoothing functions $\{\phi_k\}$ and a sequence of step lengths

Initialize iteration count $i = 0$
repeat
    Set $i \leftarrow i + 1$, $\nu \leftarrow \nu^i$, initial guess $= x^{i-1}$
    Solve sequence of smooth approximations (APP$_\nu$) using
      projected stochastic subgradient to obtain solution x^i
    Estimate risk level α^i of solution x^i
until $\alpha^i \leq \alpha_{low}$

Choosing key algorithmic parameters

Specifying an initial guess

- Determine point x^0 and bound ν^1 by solving a small-sample scenario approximation problem (to global optimality)

$$
\begin{aligned}
(\nu^1, x^0): \min_{x \in X} \; & f(x) \\
\text{s.t. } & g(x, \xi^j) \leq 0, \quad \forall j \in \{1, \dots, N\}
\end{aligned}
$$

Choosing key algorithmic parameters

Specifying an initial guess

- Determine point x^0 and bound ν^1 by solving a small-sample scenario approximation problem (to global optimality)

$$(\nu^1, x^0) : \min_{x \in X} \ f(x) \quad \text{s.t. } g(x, \xi^j) \leq 0, \quad \forall j \in \{1, \dots, N\}$$

Scaling the sequence of smoothing functions $\{\phi_k\}$

- Set $\phi_{k,j}(y) := \dfrac{1}{1 + \exp\left(\frac{-y}{\tau_{k,j}}\right)}$ for smoothing parameter $\tau_{k,j}$
- Draw $\{\xi^j\}_{j=1}^N$, set $\tau_{k,j} := O((0.1)^{k-1})\, \mathrm{median}\big(\{|g_j(x^0, \xi^j)|\}_{j=1}^N\big)$

Choosing key algorithmic parameters

Specifying an initial guess

- Determine point x^0 and bound ν^1 by solving a small-sample scenario approximation problem (to global optimality)

$$(\nu^1, x^0) : \min_{x \in X} \; f(x)$$
$$\text{s.t. } g(x, \xi^j) \leq 0, \quad \forall j \in \{1, \dots, N\}$$

Scaling the sequence of smoothing functions $\{\phi_k\}$

- Set $\phi_{k,j}(y) := \dfrac{1}{1 + \exp\left(\frac{-y}{\tau_{k,j}}\right)}$ for smoothing parameter $\tau_{k,j}$
- Draw $\{\xi^j\}_{j=1}^N$, set $\tau_{k,j} := O((0.1)^{k-1})\,\mathrm{median}\big(\{|g_j(x^0, \xi^j)|\}_{j=1}^N\big)$

Choosing step lengths

- Can leverage adaptive subgradient methods to tune step lengths

Flavor of theoretical results

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

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

Flavor of theoretical results

$$\min_{x \in X} \mathbb{E}\left[\max\left[\mathbb{1}\left[g(x,\xi)\right]\right]\right] \quad \text{(SP)}$$
$$\text{s.t. } f(x) \le \nu$$

$$\min_{x \in X} \mathbb{E}\left[\max\left[\phi_k\left(g(x,\xi)\right)\right]\right] \quad (\text{APP}_k)$$
$$\text{s.t. } f(x) \le \nu$$

Informal Theorem (Convergence of global solutions)

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

Flavor of theoretical results

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

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

Informal Theorem (Convergence of global solutions)

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

Informal Theorem (Convergence of stationary solutions)

Suppose the distribution function $F(x,\eta) := \mathbb{P}\{\max[g(x,\xi)] \leq \eta\}$ is continuously differentiable on $X \times (-\bar{\eta}, \bar{\eta})$ for some $\bar{\eta} > 0$. Then every limit point of a sequence of stationary solutions to the approximations (APP_ν) is a stationary solution to Problem (SP)

Computational Results: Portfolio Optimization

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

where $X := \left\{x \in \mathbb{R}^{1000}_{+} : \sum_i x_i = 1\right\}$, $\xi \sim \mathcal{N}(\mu, \Sigma)$

Computational Results: Portfolio Optimization

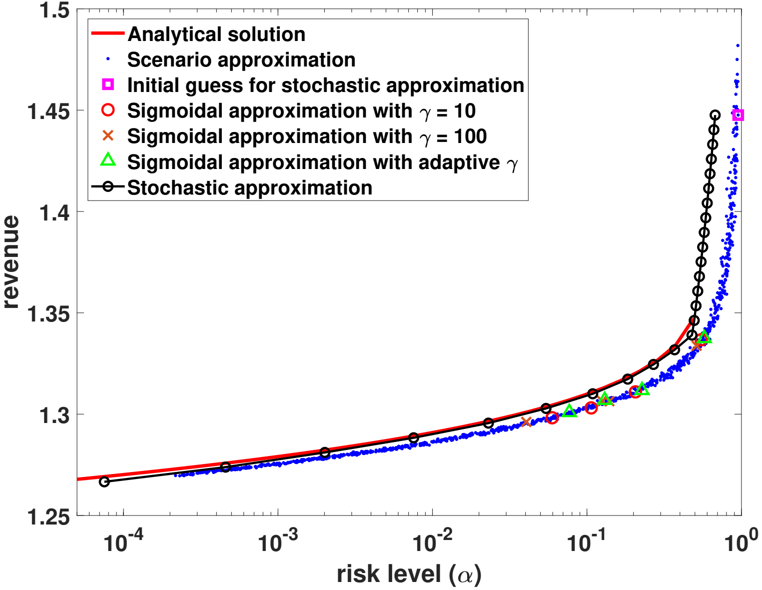
$$\max_{t,\, x \in X} \quad t$$
$$\text{s.t. } \mathbb{P}\left\{\xi^{\mathsf{T}} x \geq t\right\} \geq 1 - \alpha,$$

where $X := \left\{x \in \mathbb{R}^{1000}_{+} : \sum_i x_i = 1\right\}$, $\xi \sim \mathcal{N}(\mu, \Sigma)$

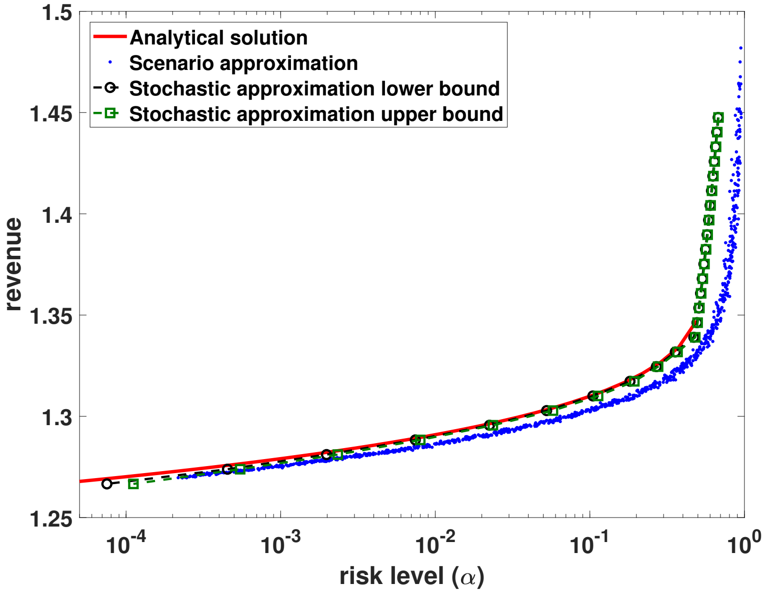
Compare

- Exact solution (deterministic equivalent)
- **Stochastic approximation** (our approach)
- Scenario approximation with sample sizes $N \in [10, 10^6]$
- Sigmoidal approximation of Cao and Zavala (2017) for risk level $\alpha = 0.01$ and with the smoothing parameters tuned
  - The sigmoidal approximation of $\mathbb{1}[\cdot]$ used by Cao and Zavala (2017) is different than our own

Computational Results: Portfolio Optimization



Results over ten different replicates



Computational Results: Norm Optimization

$$\min_{x \in \mathbb{R}^{100}_{+}} \quad -\sum_{i=1}^{100} x_i$$

$$\text{s.t.} \;\; \mathbb{P}\left\{\sum_i \xi_{ij}^2 x_i^2 \leq U^2, \;\; \forall j \in \{1,\dots,100\}\right\} \geq 1-\alpha,$$

where $\xi_{ij} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu,\sigma)$

Computational Results: Norm Optimization

$$\min_{x \in \mathbb{R}^{100}_+} \; -\sum_{i=1}^{100} x_i$$

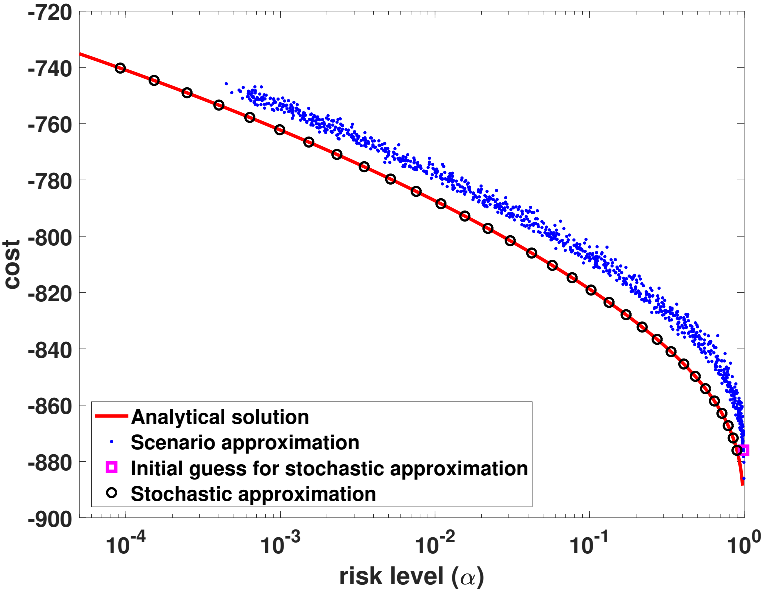
$$\text{s.t. } \mathbb{P}\left\{\sum_i \xi_{ij}^2 x_i^2 \leq U^2, \;\; \forall j \in \{1, \dots, 100\}\right\} \geq 1 - \alpha,$$

where $\xi_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma)$

Compare

- Exact solution (deterministic equivalent)
- **Stochastic approximation** (our approach)
- Scenario approximation with sample sizes $N \in [10, 10^5]$

Computational Results: Norm Optimization



Computational Results: Resource Allocation

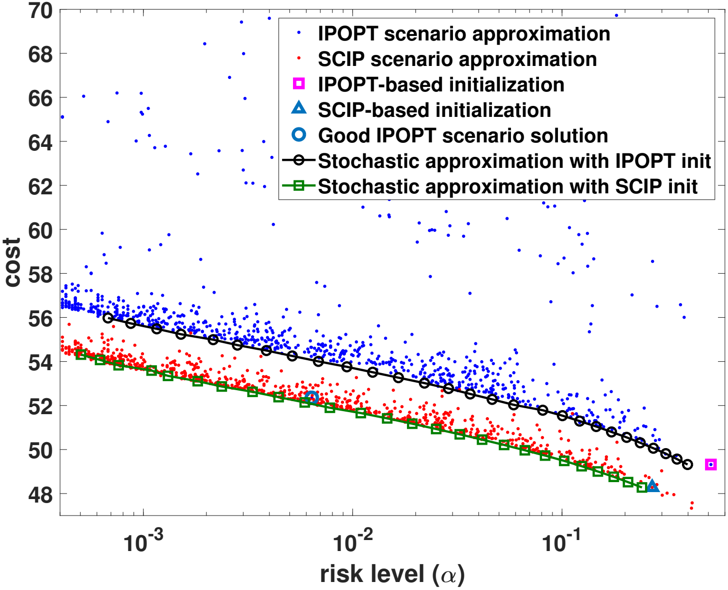
$$\min_{x \in \mathbb{R}^{20}_+} \; c^{\mathsf{T}} x$$
$$\text{s.t.} \quad \mathbb{P}\left\{x \in R(\lambda, \rho)\right\} \geq 1 - \alpha,$$

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, $\quad c_i$: 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

Computational Results: Resource Allocation



Concluding Remarks

Proposed a stochastic subgradient method for approximating the efficient frontier of chance-constrained NLPs

- Efficient frontier can help make informed decisions
- Smoothing + implicit sampling helps avoid bad local minima
- Harness the power of stochastic subgradient methods
- Consistently outperforms existing approaches

Concluding Remarks

Proposed a stochastic subgradient method for approximating the efficient frontier of chance-constrained NLPs

- Efficient frontier can help make informed decisions
- Smoothing + implicit sampling helps avoid bad local minima
- Harness the power of stochastic subgradient methods
- Consistently outperforms existing approaches

Paper: Math. Programming Computation, 13(4), pp. 705-751
Code: `https://github.com/rohitkannan/SA-for-CCP`

Concluding Remarks

Proposed a stochastic subgradient method for approximating the efficient frontier of chance-constrained NLPs

- Efficient frontier can help make informed decisions
- Smoothing + implicit sampling helps avoid bad local minima
- Harness the power of stochastic subgradient methods
- Consistently outperforms existing approaches

Paper: Math. Programming Computation, 13(4), pp. 705-751
Code: `https://github.com/rohitkannan/SA-for-CCP`

Interesting research directions

- Handling deterministic nonconvex constraints directly
- Reducing effort spent on projections
- Extension to distributionally robust chance constraints

Questions? rohitk@alum.mit.edu

References I

L. Andrieu, G. Cohen, and F. Vázquez-Abad. Stochastic programming with probability. *arXiv preprint arXiv:0708.0281*, 2007.

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

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

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

A. Charnes and W. W. Cooper. Chance-constrained programming. *Management Science*, 6(1):73–79, 1959.

A. Charnes, W. W. Cooper, and G. H. Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263, 1958.

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

D. Davis and D. 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

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

C. Gotzes, H. Heitsch, R. Henrion, and R. Schultz. On the quantification of nomination feasibility in stationary gas networks with random load. *Mathematical Methods of Operations Research*, 84(2):427–457, 2016.

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

C. M. Lagoa, X. Li, and M. Sznaier. Probabilistically constrained linear programs and risk-adjusted controller design. *SIAM Journal on Optimization*, 15(3):938–951, 2005.

R. Lepp. Extremum problems with probability functions: Kernel-type solution methods. 2009.

P. Li, H. Arellano-Garcia, and G. Wozny. Chance constrained programming approach to process optimization under uncertainty. *Computers & Chemical Engineering*, 32 (1-2):25–45, 2008.

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

References III

J. Luedtke and S. Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008.

B. L. Miller and H. M. Wagner. Chance constrained programming with joint constraints. *Operations Research*, 13(6):930–945, 1965.

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

V. I. Norkin. The analysis and optimization of probability functions. 1993.

A. Peña-Ordieres, J. R. Luedtke, and A. Wächter. Solving chance-constrained problems via a smooth sample-based nonlinear approximation. *SIAM Journal on Optimization*, 30(3):2221–2250, 2020.

A. Prékopa. On probabilistic constrained programming. In *Proceedings of the Princeton symposium on mathematical programming*, pages 113–138. Princeton, NJ, 1970.

A. Prékopa. *Stochastic programming*, volume 324. Springer Science & Business Media, 1995.

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

L. Roald, F. Oldewurtel, T. Krause, and G. Andersson. Analytical reformulation of security constrained optimal power flow with probabilistic constraints. In *2013 IEEE Grenoble Conference*, pages 1–6. IEEE, 2013.

References IV

R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42, 2000.

A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

W. van Ackooij and R. Henrion. Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. *SIAM Journal on Optimization*, 24(4):1864–1889, 2014.

W. van Ackooij and R. Henrion. (Sub-)Gradient formulae for probability functions of random inequality systems under Gaussian distribution. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):63–87, 2017.