Importance sampling for chance-constrained optimization

Vyacheslav Vadimovich Gorchakov

MIPT

Курс: Автоматизация научных исследований
(практика, В. В. Стрижов)/Группа 874

Эксперт: Y. . Maximov

Консультант: Ю. . Maximov

2021

Project's goal

Time-efficient and reliable solution to the (linear) chance-constrained optimization problem

$$\begin{aligned} &\min_{x} \texttt{cost}(x) \\ \texttt{s.t.:}\, &\mathbb{P}_{\xi}(A(x+\xi) \leq b) \geq 1-\eta, \eta \to 0+ \\ &\xi \sim \mathcal{N}(\mu, \Sigma), \end{aligned}$$

where $\texttt{cost}(x)$ is convex in x.

Applications

- Risk-aware portfolio optimization;
- Network reliability assessment;
- Power generation cost minimization with security constraints

Challenge and state-of-the-art

Where is a challenge?

- The problem is NP-hard and non-convex (Khachyan'1979)
- Existing relaxations overestimate the failure probability, does not match the exact solution even in $\mathbb{R}^1$

Convex approximations:

- Substitute the problem by

$$\min_{x, t>0} f(x)$$
$$\texttt{s.t.:}\, t^{-1}\mathbb{E}_{\xi}(\psi(A(x+\xi)/t)) \geq \eta$$
$$\xi \in \mathcal{N}(\mu, \Sigma)$$

  where ψ is any convex function that upperbounds the indicator
- The constraints becomes a deterministic convex constraint.

State-of-the-art approximations are:

- the Markov's approximation
- the Chebyshev's approximation
- the Bernstein's approximation.

Relaxations and Scenario Approximation

Popular relaxations are:

- $\psi(u) = (1+u)_+$ for the Markov's approximation
- $\psi(u) = (1+u)_+^2$ for the Chebyshev's approximation
- $\psi(u) = e^u$ for the Bernstein's approximation.

Sample Average Approximation (SAA)

Sample average approximation

- theoretically solves the problem, especially in cases $\eta \approx 1$
- requires a lot of samples.
- the main idea is to replace the probability constraints by sampling.

$$c^\top x \to \min_{x \in \mathbb{R}^n}$$

$$\text{s.t.:} \frac{1}{N} \sum_{i=1}^{N} I_{(0,+\infty)}(g(x,\xi_i)) \le \gamma$$

Case $\gamma > \eta$ corresponds to the solution, which is a lower bound on the original problem. Case $\gamma = 0$ is called the scenario approximation.

Publications

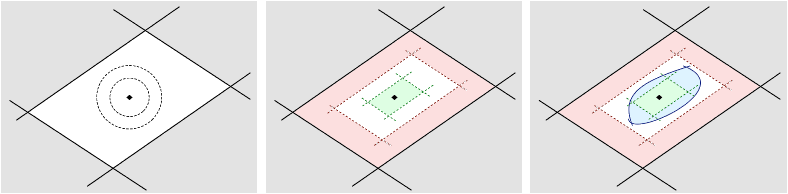
1. Feller, William An introduction to probability theory and its applications.
2. Laguel, Yassine and Van Ackooij, Wim and Malick, Jérôme and Ramalho, Guilherme On the Convexity of Level-sets of Probability Functions
3. Laguel, Yassine and Malick, Jérôme and Ackooij, Wim Chance constrained problems: a bilevel convex optimization perspective
4. Karimi, Hamed and Nutini, Julie and Schmidt, Mark Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition

Algorithm IMPLICATION: Importance Sampling for Linear CC-Optimization

Require: Optimization problem, Eq. **??**
Ensure: (ε) optimal solution, e.g. point x^*, so that

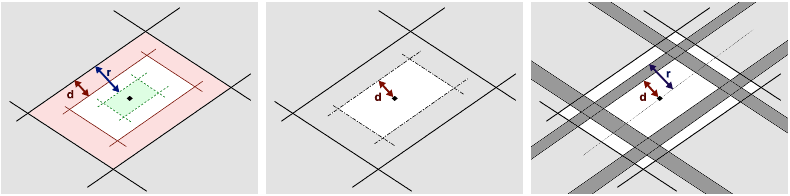
```
y ⇐ 1
if n < 0 then
   X ⇐ 1/x, N ⇐ −n
else
   X ⇐ x, N ⇐ n
end if
while N ≠ 0 do
   if N is even then
      X ⇐ X × X, E N ⇐ N/2
   else {N is odd}
      y ⇐ y × X, N ⇐ N − 1
   end if
end while
```

Слайд с визуализацией из CC-opt



- Left: feasibility set $\mathcal{P} = \{x : \; a_i^\top x \le b_i, 1 \le i \le m\}$. Dashed lines – level sets of Gaussian distribution. Black diamond represent an optimization variable x.
- Middle: all points in green area satisfy $\sum_{i=1}^m \mathbb{P}_\xi(a_i^\top(x+\xi) > b_i) \le \eta$.
- Right: Feasibility area boundary marked as blue. All points in green and blue satisfy $\mathbb{P}_\xi\{\bigcap_{i=1}^m a_i^\top(x+\xi) \le b_i\} \ge 1-\eta$.

Слайд с визуализацией из CC-opt



- Left: d and r are the distances between the polytope boundary and boundaries of the green and the red zones.
- Middle: Sampling in white area does not contribute to the error of scenario approximation.
- Right: More accurately, we are interested in samples only in dark-grey area. Samples that closer diamond are always feasible and samples that are far from diamond are always infeasible.

Experiments

First experiment

It is devoted to the simplest case with linear function $g(x, \xi)$:

$$c^\top x \to \min$$
$$(a^\top(x + \xi) \leq b) \geq 1 - \eta,$$

where $a \in \mathbb{R}^n$ is a fixed vector, x is a an optimization variable, and $\xi \in \mathcal{N}(\mu, \Sigma)$.

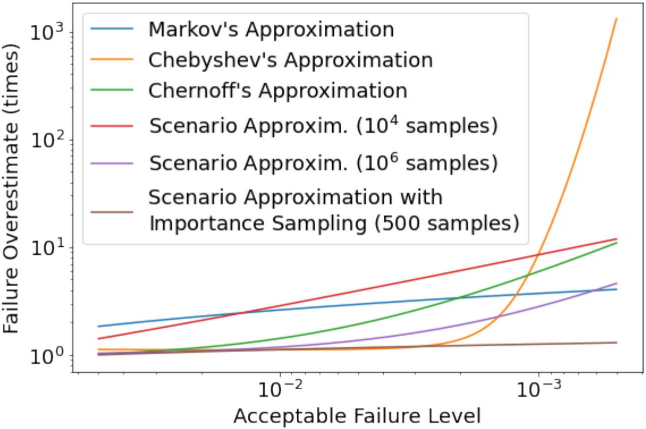
Second experiment

DC grid, and DC Optimal power flow:

$$c^\top p \to \min$$
$$p = B\theta, p_i^{\max} \geq p_i \geq p_i^{\min}, \theta_i^{\max} \geq \theta \geq \theta_i^{\min}$$
$$p \sim \mathcal{N}(\mu, \Sigma),$$

where $\Sigma = \mathrm{Diag}(\Sigma)$, $\sqrt{\Sigma_{ii}} = 0.05 p_i$.

Approximations performance comparison



- $\psi(u) = (1+u)_+$ for the Markov's approximation
- $\psi(u) = (1+u)_+^2$ for the Chebyshev's approximation
- $\psi(u) = e^u$ for the Bernstein's approximation.

Conclusions

- Propose scenario approximation to DC Optimal power flow problem
- Scenario Approximation with Importance Sampling gives a more accurate result than state-of-the-art approximations, especially in cases $\eta \approx 1$