Importance Sampling for Chance Constrained Optimization*

Vyacheslav Gorchakov †

Abstract. Stochastic optimization problems with probabilistic constraints have a wide-range of applications in engineering. In particular, chance constrained optimization appears in many flavors of engineering, where impose an extra condition of a system reliability under random noise. Power systems is a prominent example, where one needs to minimize energy cost generation; however, with a high probability the system should be within the security limits under random fluctuations of generation/demand.The problem is NP-hard even in the simplest cases. Although several approximations are known to be quite successful. A scenario approximation, which consists in sampling random points according to the (nominal) noise distribution and solving a deterministic problem afterwards, is known to be one of the most efficient. In our study, we propose a way to reduce the complexity of scenario approximation for the simplest case of linear objective and linear chance constraints with Gaussian noise by utilizing importance sampling. This case is popular in several applications, including DC approximation.

Key words. chance constrained optimization, importance sampling

AMS subject classifications. 68Q25, 68R10, 68U05

1. Introduction.

Chance constrained optimization methods have many applications in computer vision, robotics, power systems, self-driving, gas transmission systems, water and heating networks and transportation networks. The origins of such methods date back to desicion theory in 1950 and, more recently, to robust optimization methods, where we are dealing with probabilistic and non-probabilistic models of robustness. Since a robust solution is usually conservative, it is desired to have a compromise between the robustness and the risk of failure in many practical problems. Therefore, chance-constrained optimization approaches are more often proposed in engineering practice.

Firstly, it must be noted that in general case finding exact solution of optimization problem is NP-complete problem. This result was achieved by A. Khatchyan in 1978. The convexity of constraints simplifies chance constrained problems and there are some state-of-the art approaches(CVAR, Bernstein approximation, SAA) to find a convex set to approximate the feasible set. Bernstein approximation offers to find a convex set inside the feasible set, but Sample average approximation is a good approach especially, when the confidence level η is approximately 1, but requires a lot of samples.

Sampling from existing Gaussian distribution generates points mostly far from the border of sample set. Importance sampling allows to avoid this problem, since our goal is to consider points near the set border.

The proposed approach is to use importance sampling with sample average approximation. Probably, it requires less quantity of samples. We will try to empirically compare various approaches, theoretically and contribute to the sample complexity. We will work with

*Submitted to the editors DATE.

Funding: This work was funded by the Fog Research Institute under contract no. FRI-454.

† (ddoe@imag.com, http://www.imag.com/~ddoe/).

synthetic data from Matpower (IEEE cases in power systems). Our goal in this paper is to reduce sample complexity of scenario approximation and figure out which of the state of the art approaches performs better, and identify the cases where SAA with importance sampling is the best choice.

The paper is organized as follows. Our main results are in **??**, our new algorithm is in **??**, experimental results are in section 6, and the conclusions follow in section 8.

2. Background and Problem Setup.

Let $\{x_i\}_{i=1}^m, x_i \in \mathbb{R}^n$ be a sequence of points. Let $f(x) = c^\top x$ be a linear function. We consider linear probability constraints:

$$\mathrm{Prob}(g(x,\xi) \geq 0) \leq \eta$$

where ξ is scenario, η is "confidence" level, $g(\cdot,\cdot)$ is a linear function of x, ξ in a simple approach. If successful, the approach will be extended to non-linear models.

So, let state optimization problem:

$$\begin{aligned} & c^\top x \to \min_{x\in\mathbb{R}^n} \\ & \text{s.t.: } \mathrm{Prob}_\xi(A(x+\xi) \geq b) \leq 1-\epsilon \\ & \quad \xi \sim \mathcal{N}(\mu, \Sigma) \\ & \quad \epsilon \ll 1 (\text{eventual convexity holds}) \end{aligned}$$

As a loss function, in our study we will use hinge loss, quadratic loss, soft-max functions. We will compare results of using these functions on syntetic data.

Firstly, we notice that probability constraints can be approximated:

$$Prob(g(x,\xi) \geq 0) \leq \frac{1}{n}\sum_{i=1}^{n} h_p(x,\xi) + O(\frac{Var}{\sqrt{n}}).$$

Let $\{\xi_i\}_{i=1}^n$ be sequence of scenarios. Let $g(x,\xi) = A(x+\xi_i)$. If we have hinge loss function with parameter p, then:

$$\mathbb{E}A(x+\xi_i) \leq \sum_{i=1} h_p(A(x+\xi_i) \leq b) + O(N^{-0.5})$$

As state-of-the-art methods it is known such convex approximations of feasible set as Bernshtein approximation, Markov approximation, Tchebyshev approximation and SAA (sample average approximation).

Bernshtein approximation

Here we assume a surrogate function $\phi(u) = e^u$ to replace and approximate the 0-1 loss function. Then constraints, which we set up becomes:

$$t \log \mathbb{E}[e^{g(x,\xi)/t}] - t\log\eta \leq 0$$

And after transformations, which are detailed in [1], we have:

$$t(\sum_{i=1}^{n} \log \mathbb{E}[e^{\xi x_i/t}]) - t \log \eta \leq 0$$

The constraints becomes a deterministic convex constraint.

Sample average approximation Sample average approximation theoretically solves the problem, especially in cases $\eta \approx 1$, but requires a lot of samples.

The main idea in this approach is to replace the probability constraints by sampling.

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

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

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

Primal-dual stochastic optimization http://www.cs.toronto.edu/~slwang/primal-dual.pdf, http://webdoc.sub.gwdg.de/ebook/serien/e/CORE/dp2005_67.pdf, https://projecteuclid.org/journals/stochastic-systems/volume-4/issue-1/Deterministic-and-stochastic-primal-dual-subgradient-algorithms-f 10.1214/10-SSY010.full

3. "Main results".

4. Empirical Study.

5. Conclusion.

Summary + Extensions + Applications

cross-validation procedure, restrictions to the solutions, external (industrial) quality criteria,

cross-validation – once we have received the "optimal point" we can check the probability of breaking the constraints

external quality criteria – at the "optimal" point check the probability of violating the constraints by sampling – should be $1 - \varepsilon$ or about. If more – the solution is very conservative, if less – a bug in the algorithm and/or implementation.

$Prob(g(x, \xi) \geq 0) = \int f(x)\mu(x)dx = \int \frac{f(x)\mu(x)\eta(x)}{\eta(x)}dx$.

This integral can be approximated: $\int \frac{f(x)\mu(x)\eta(x)}{\eta(x)}dx = \frac{1}{N} \sum\limits_{i} f(x_i)\mu(x_i)$.

Then $Var(\frac{1}{N} \sum\limits_{i} f(x_i)\mu(x_i) = \frac{1}{N} \int \frac{f^2(x)\mu^2(x)}{\eta(x)}dx - \frac{P^2}{N}$.

6. Experimental results.

Quisque facilisis auctor sapien. Pellentesque gravida hendrerit lectus. Mauris rutrum sodales sapien. Fusce hendrerit sem vel lorem. Integer pellentesque massa vel augue. Integer elit tortor, feugiat quis, sagittis et, ornare non, lacus. Vestibulum posuere pellentesque eros. Quisque venenatis ipsum dictum nulla. Aliquam quis quam non metus eleifend interdum. Nam eget sapien ac mauris malesuada adipiscing. Etiam eleifend neque sed quam. Nulla facilisi. Proin a ligula. Sed id dui eu nibh egestas tincidunt. Suspendisse arcu.

Figure 1 shows some example results. Additional results are available in the supplement in Table SM1.

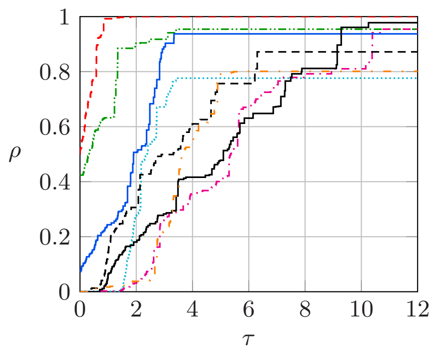

Figure 1. *Example figure using external image files.*

Maecenas dui. Aliquam volutpat auctor lorem. Cras placerat est vitae lectus. Curabitur massa lectus, rutrum euismod, dignissim ut, dapibus a, odio. Ut eros erat, vulputate ut, interdum non, porta eu, erat. Cras fermentum, felis in porta congue, velit leo facilisis odio, vitae consectetuer lorem quam vitae orci. Sed ultrices, pede eu placerat auctor, ante ligula rutrum tellus, vel posuere nibh lacus nec nibh. Maecenas laoreet dolor at enim. Donec molestie dolor nec metus. Vestibulum libero. Sed quis erat. Sed tristique. Duis pede leo, fermentum quis, consectetuer eget, vulputate sit amet, erat.

7. Discussion of $Z = X \cup Y$. Curabitur nunc magna, posuere eget, venenatis eu, vehicula ac, velit. Aenean ornare, massa a accumsan pulvinar, quam lorem laoreet purus, eu sodales magna risus molestie lorem. Nunc erat velit, hendrerit quis, malesuada ut, aliquam vitae, wisi. Sed posuere. Suspendisse ipsum arcu, scelerisque nec, aliquam eu, molestie tincidunt, justo. Phasellus iaculis. Sed posuere lorem non ipsum. Pellentesque dapibus. Suspendisse quam libero, laoreet a, tincidunt eget, consequat at, est. Nullam ut lectus non enim consequat facilisis. Mauris leo. Quisque pede ligula, auctor vel, pellentesque vel, posuere id, turpis. Cras ipsum sem, cursus et, facilisis ut, tempus euismod, quam. Suspendisse tristique dolor eu orci. Mauris mattis. Aenean semper. Vivamus tortor magna, facilisis id, varius mattis, hendrerit in, justo. Integer purus.

8. Conclusions. Some conclusions here.

Appendix A. An example appendix. Aenean tincidunt laoreet dui. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Integer ipsum lectus, fermentum ac, malesuada in, eleifend ut, lorem. Vivamus ipsum turpis, elementum vel, hendrerit ut, semper at, metus. Vivamus sapien tortor, eleifend id, dapibus in, egestas et, pede. Pellentesque faucibus. Praesent lorem neque, dignissim in, facilisis nec, hendrerit vel, odio. Nam at diam ac neque aliquet viverra. Morbi dapibus ligula sagittis magna. In lobortis. Donec aliquet ultricies libero. Nunc dictum vulputate purus. Morbi varius. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In tempor. Phasellus commodo porttitor magna. Curabitur vehicula odio vel dolor.

Lemma A.1. *Test Lemma.*

Acknowledgments. We would like to acknowledge the assistance of volunteers in putting together this example manuscript and supplement.

REFERENCES