Extended Abstract
The 55th ISCIE International Symposium on Stochastic Systems Theory and Its Applications, Tokyo, Japan, Nov. 17-18, 2023

Risk Assessment for Sparse Optimization with Relaxation

Zhicheng Zhang† and Yasumasa Fujisaki†

†Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka 565-0871, Japan
E-mail: {zhicheng-zhang, fujisaki}@ist.osaka-u.ac.jp

Sparse optimization is a widely accepted methodology that allows one to generate sparse solutions by penalizing an exact ℓ_0 norm on the decision variables [1]. Meanwhile, a chance constrained representation for uncertain constraints is a common strategy to assess the risk level for the obtained feasible solutions. Based on the above setups, a *chance constrained sparse optimization problem* is well-defined that can not only measure the sparse cost but also evaluate the risk of constraints violation. To make such program computationally tractable, we make a convex relaxation for the ℓ_0 cost and approximate the chance constraints by randomly taking a data-driven sampling for uncertain parameters [2], which can reduce to a random convex program [3]. In this context, we focus on how to *make a trade-off bridge between the sparse cost and the risk level by relaxing the constraint violations.* We then shift the idea from a relaxed sparse convex optimization to risk-aware sparse optimal control application.

Let us start from a *chance constrained sparse optimization problem* ($\mathrm{CCSOP}^0_\epsilon$), defined as follows[1]

$$\mathrm{CCSOP}^0_\epsilon: \quad \begin{array}{ll} \min\limits_{x\in\mathbb{X}} & \|x\|_0 \\ \text{s.t.} & \mathbb{P}\left\{q\in\mathbb{Q} : h(x,q)\le 0\right\} \ge 1-\epsilon, \end{array} \tag{1}$$

where $\epsilon \in (0,1)$ represents the risk (or constraint violation) level for the chance constrained framework.

Definition 1 (Risk) *Given a decision variable x, the probability of violation (or risk) is defined as*

$$V(x) = \mathbb{P}\big\{q\in\mathbb{Q} : \ h(x,q) > 0\big\}.$$

Hence, $V(x)$ is the probability with which the violation constraint for $\mathrm{CCSOP}^0_\epsilon$ (1). We say that x has an *ϵ-probabilistic robustness* if it holds that $V(x)\le\epsilon$.

It is noticed that the proposed $\mathrm{CCSOP}^0_\epsilon$ (1) is a non-convex program. Because of the fact that the *non-convex* and *non-smooth* nature of the objective (or cost) function, which precisely measures the sparsity $\|u\|_0$. On the other hand, the feasibility of the program (1) involves the evaluation of multiple integrals for risk (see Definition 1), which is often computationally intractable.

Instead of an exact sparsity-inducing ℓ_0 norm for $\|x\|_0$, we make a convex relaxation, namely, take ℓ_1 cost $\|x\|_1$ in program (1). Accordingly, the uncertain parameter q is modeled as a random outcome that *independent and identically distributed* (i.i.d.) draws N random "samples" or "scenarios" (i.e., $q^1, q^2, \ldots, q^N$) from some probability $\mathbb{P}$ [2].

To make the above operations more concrete, one can consider the following ℓ_1 relaxed *scenario-based sparse convex optimization problem* (SSCOP^1_N), formulated as

$$\mathrm{SSCOP}^1_N: \quad \begin{array}{ll} \min\limits_{x\in\mathbb{X}} & \|x\|_1 \\ \text{s.t.} & h(x,q^i)\le 0, \quad i=1,\ldots,N. \end{array} \tag{2}$$

The solution to program (2), denoted by x^*_N, is a *data-driven, randomized, sparse solution* based on the N observations or scenarios. In other words, SSCOP^1_N is a random convex program [3].

Assumption 1 (Existence and Uniqueness) *For every N and for every sample q^i, $i=1,\ldots,N$, there exists a solution of the program (2), which becomes unique after the application of a tie-break rule.*

[1]Notation: Let $\mathbb{X}\subseteq\mathbb{R}^n$ be a compact convex set and $(\mathcal{Q},\mathfrak{B}(\mathcal{Q}),\mathbb{P})$ be a probability space, where $\mathcal{Q}$ is a metric space with respect to Borel σ-algebra $\mathfrak{B}(\mathcal{Q})$. A measurable *uncertain function* $h:\mathbb{X}\times\mathcal{Q}\to\mathbb{R}$, which is *convex* in the first argument x for each $q\in\mathcal{Q}$, and *bounded* in the second argument q for each $x\in\mathbb{X}$. Denote $\|x\|_0$ as the ℓ_0 quasi-norm of the vector $x\in\mathbb{R}^n$, which depends on its support $\|x\|_0=\mathrm{supp}\{i : x_i\ne 0\}$ that counts the number of nonzero elements to measure its sparsity (resp., a convex relaxation of ℓ_0 norm of the vector x is defined as ℓ_1 norm $\|x\|_1$ that sums of its absolute values of the components $\|x\|_1=\sum_i|x_i|$, which still induces the sparsity).

Assumption 2 (Non-accumulation) *For every* $x \in \mathbb{X}$, $\mathbb{P}\{q : h(x,q) = 0\} = 0$.

An intriguing and profound question arises: How can one effectively balance the *sparse cost* and the associated *risk* in program (2)? Motivated by the scenario program having a linear objective function [4, Sec. 4], it suggests us exploring a flexible paradigm that performs a more "soft" constraints for program (2). In this context, we slightly modify the program as sparse program, called *sparse optimization with relaxation*, which gives rise to a *relaxed* scenario sparse optimization problem (SSCOP_N^ρ), described by

$$\mathrm{SSCOP}_N^\rho: \quad \begin{aligned} \min_{x\in\mathbb{X},\xi^i\geq 0} \quad & \|x\|_1 + \rho\sum_{i=1}^N \xi^i \\ \text{s.t.} \quad & h(x,q^i) \leq \xi^i, \quad i = 1,\ldots,N, \end{aligned} \tag{3}$$

where $q^1, q^2, \ldots, q^N$ are i.i.d. random samples from probability space $(\mathcal{Q}^N, \mathfrak{B}^N(\mathcal{Q}), \mathbb{P}^N)$, and $\xi^i \geq 0$ are the "relaxed" slack variables that indicate the maximum constraint violation over all possible uncertain parameters q. The program (3) has $n+N$ decision variables, including $x \in \mathbb{R}^n$, $\xi = [\xi^1, \ldots, \xi^N]^\top \in \mathbb{R}^N$. When $\xi^i > 0$, the constraint $h(x,q^i) \leq 0$ in (2) is here relaxed to $h(x,q^i) \leq \xi^i$ that results in a "regret" ξ^i. Meanwhile, the penalty weight ρ is used to make a trade-off between minimizing the sparse cost and the regrets for constraint violations. For large enough ρ value, for example $\rho \to \infty$, it goes back to SSCOP_N^1 (2).

It is well known that optimal control problem with constraints relaxation plays a vital role in control design, since the decision-makers have to simultaneously take the performance and cost into account in real-world applications. In what follows, we provide the following problem taken from sparse robust control design [3], which is a main focus of attention in the present context.

Problem 1 (Risk-aware Sparse Optimal Control Problem - RaSOCP) *Given a discrete reachable uncertain linear time invariant (LTI) system* $z_{k+1} = A(q)z_k + B(q)u_k$, *where* $u_k \in \mathbb{R}^m$ *is the input,* $z_k \in \mathbb{R}^n$ *is the corresponding state, and uncertain system matrices* $A(q) \in \mathbb{R}^{n\times n}, B(q) \in \mathbb{R}^{n\times m}$ *depend on the uncertainty* $q \in \mathbb{Q} \subseteq \mathbb{R}^{n_q}$. *For given parameters* z_0, L, ρ, γ, *one aims to seek a control sequence* $\{u_k\}_{k=0}^{L-1}$ *with input sparsity [1] such that it drives the state near to the target* $\bar{z} \in \mathbb{R}^n$ *(e.g., the origin* $\bar{z} = 0$*) as well as hedges against the uncertainty* $\{q^i\}_{i=1}^N$ *[2]. Realizing trade-off between the sparse control input and the risk assessment amounts to solving a revised program* SSCOP_N^ρ, *defined as follows*

$$\mathrm{RaSOCP}_N^\rho: \quad \begin{aligned} \min_{u\in\mathbb{R}^{mL},\xi^i\geq 0} \quad & \|u\|_1 + \rho\sum_{i=1}^N \xi^i \\ \text{s.t.} \quad & h(u,q^i) \leq \xi^i, \quad i = 1,\cdots,N. \end{aligned} \tag{4}$$

Here the scalar-valued uncertain constraint $h(u,q)$ *is measured by a distance between the terminal state* $z_L(u,q)$ *and the target* $\bar{z}$ *with a metric* $\gamma \geq 0$, *defined by*

$$h(u,q) \doteq \|z_L(u,q) - \bar{z}\|_2 - \gamma = \|A(q)^L z_0 + \mathfrak{R}(q)u - \bar{z}\|_2 - \gamma \tag{5}$$

with $\mathfrak{R}(q) = \begin{bmatrix} A(q)^{L-1}B(q) & \ldots & A(q)B(q) & B(q) \end{bmatrix}, \quad u = \begin{bmatrix} u^\top(0) & u^\top(1) & \ldots & u^\top(L-1) \end{bmatrix}^\top.$

Based on Assumptions 1 and 2, the program (4) is computationally tractable, and a general theory for scenario (linear) program with relaxation (but not involves sparse cost) has been shown in [4, Theorem 4]. When a sparse cost $\|u\|_1$ is performed, we can interpret it as a regularization problem, which will reduce the number of nonzero elements of the decision variables, for instance $\|u\|_0 \leq s$, $s \ll mL$. It is also important to provide a *probabilistic robustness* and *trade-off* guarantees for risk-aware sparse optimal control problem.

References

[1] M. Nagahara, *Sparsity methods for systems and control.* Now Publishers, 2020.

[2] M. C. Campi and S. Garatti, *Introduction to the scenario approach.* SIAM, 2018.

[3] Z. Zhang and Y. Fujisaki, "Sparse robust control design via scenario program," in *Proceeding of the 53rd ISCIE Int. Symp. Stoch. Syst. Theor. Appli. (SSS'21)*, vol. 2022, pp. 61–64, ISCIE, 2022.

[4] S. Garatti and M. C. Campi, "Risk and complexity in scenario optimization," *Mathematical Programming*, vol. 191, no. 1, pp. 243–279, 2022.