Proceedings of the 55th ISCIE International Symposium
on Stochastic Systems Theory and Its Applications
Adachi, 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

Abstract

Sparse optimization with uncertainty is a widely accepted methodology that allows one to obtain robust feasible solution and then conduct sparse decision-making. To alleviate the conservative solution for robust counterpart, a probabilistic problem setup for uncertain parameter is employed to assess the risk level for the candidate solutions. Therefore, 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. In this context, we are interested in *making 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.

1 Introduction

Sparse optimization is one of critical issues for cost-effectiveness and recourse allocations. Needless to say, sparse modeling has been widely applied in diverse fields, including sparse support vector machine in machine learning, Lasso regularization in statistical inference, compressed sensing in signal processing, etc. Subsequently, when the studied datasets or models encountering with uncertainty, like disturbances, noise or model mismatches, robustness plays an important role to provide the high reliability and safety guarantees for decision-making under uncertainty.

In this article, we deal with the sparse optimization problem in uncertainty, where the sparsity is used to promote the less cost performance for optimizer and the uncertainty is modeled as the chance-constrained setting to reveal the resilient and soft constraints. For example, in control system applications, the engineer aims to seek a minimum control efforts for the designed controller [1] as well as ensure the system behavior would hedge against the stochastic perturbations [2], referred to as risk-aware sparse optimal control problem [3]. More precisely, "sparse control" is to minimize the length of time horizon over which the control signal is exactly zero by performing the ℓ_1 norm convex optimization. Meanwhile, the "risk-aware" evaluates the *probabilistic robustness* for the uncertain parameters by randomly i.i.d. generating a finite number of samples to return randomized scenario solutions.

Recently, a general theory called risk and complexity in scenario optimization was proposed in [4] that further explores the connection between the violation of probability (i.e., risk) and the number of support scenarios (i.e., complexity). A major difference of the current work from the method in [3] is that we propose a *posterioi* method for estimating the optimal solution. That is, the optimal solution is firstly computed through a testing data, and then provides a posteriori guarantee for the optimal solution using a finite validation data samples, as investigated in [4]; not directly gives a *prior* and *explicit* sample complexity for scenarios.

Motivated by this, we consider the sparse optimization with relaxation and its application to risk-aware sparse optimal control for a discrete-time uncertain linear time invariant system. Finally, we give the numerical example to illustrate the effectiveness of our proposed method in sparse control problem.

2 Problem Formulation

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

$$\begin{aligned} \min_{x\in\mathcal{X}} \quad & \|x\|_0 \\ \text{s.t.} \quad & \mathbb{P}\left\{q \in \mathcal{Q} : h(x,q) \le 0\right\} \ge 1-\epsilon, \end{aligned} \tag{1}$$

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 (i.e., risk) is defined as*

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

[1]Notation: Let $\mathcal{X} \subseteq \mathbb{R}^n$ 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* $h : \mathcal{X} \times \mathcal{Q} \to \mathbb{R}$, which is *convex* in the first argument x for each $q \in \mathcal{Q}$, and *bounded* in the second argument q for each $x \in \mathcal{X}$. Denote $\|x\|_0$ as the ℓ_0 quasi-norm of the vector $x \in \mathbb{R}^n$, which depends on its support $\|x\|_0 = \mathrm{supp}\{i : x_i \neq 0\}$ 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).

Hence, $V(x)$ is the probability with which the violation constraint for CCSOP_ϵ^0 (1). For $\epsilon \in (0,1)$, x attains an *ϵ-probabilistic robustness* if it holds that $V(x) \le \epsilon$.

Notice that the proposed CCSOP_ϵ^0 (1) is a non-convex optimization and NP-hard problem. On the one hand, both the objective function and chance constraints keep non-convex structure to measure the exact sparsity and risk, respectively. On the other hand, the feasibility of the probabilistic constraint in program (1) involves the calculation of multiple integrals, which is often computationally intractable.

Instead of the precise non-convex ℓ_0 norm for decision variable, 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., $q^1, q^2, \dots, q^N$) from the probability distribution $\mathbb{P}$ [2,5].

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

$$\begin{aligned} \min_{x\in\mathcal{X}} \quad & \|x\|_1 \\ \text{s.t.} \quad & h(x, q^i) \le 0, \quad i = 1, \dots, N. \end{aligned} \tag{2}$$

The solution to program (2), denoted by x_N^*, is a *data-driven, randomized, sparse solution* based on the N observations. Obviously, SSCOP_N^1 is essentially a sparse random convex program [3].

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

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

An intriguing and profound question arises: How can one effectively balance the *sparse cost performance* and the associated *violation of constraints* in program (2)? Motivated by the scenario program having a linear objective function [4, Section 4], it suggests us exploring a flexible paradigm that minimizes the cost as well as the additional "soft" constraints for program (2).

Adapted from program (2), this context focuses on risk assessment for *sparse optimization with relaxation*, giving rise to a *relaxed* scenario-based sparse convex optimization problem (SSCOP_N^ρ), described by

$$\begin{aligned} \min_{x\in\mathcal{X}, \xi^i \ge 0} \quad & \|x\|_1 + \rho \sum_{i=1}^N \xi^i \\ \text{s.t.} \quad & h(x, q^i) \le \xi^i, \quad i = 1, \dots, N, \end{aligned} \tag{3}$$

where $q^1, q^2, \dots, q^N$ are i.i.d. random samples from probability space $(\mathcal{Q}^N, \mathfrak{B}^N(\mathcal{Q}), \mathbb{P}^N)$, and $\xi^i \ge 0$ are the "relaxed" slack variables that indicate the maximum constraint violation over all possible uncertain parameters q^i, and the allowed violations ξ^i's are penalized to the original cost function.

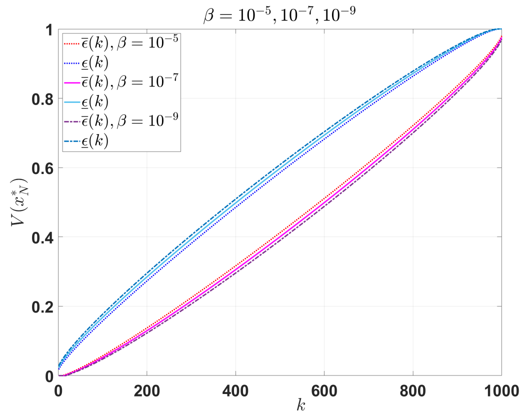

Fig. 1: Profile of curves $\underline{\epsilon}(k)$ and $\overline{\epsilon}(k)$ for the violation of probability $V(x_N^*)$ with $N = 1000$, and $\beta = 10^{-5}, 10^{-7}, 10^{-9}$, respectively, where $k = 0, 1, \cdots, N$.

Remark 1 *This program* (3) *has $n + N$ decision variables, including $x \in \mathbb{R}^n$ and $\xi = [\xi^1, \dots, \xi^N]^\top \in \mathbb{R}^N$. When $\xi^i > 0$, the constraint $h(x, q^i) \le 0$ in* (2) *is here relaxed to $h(x, q^i) \le \xi^i$ that results in a "regret" ξ^i. Meanwhile, the penalty weight ρ is used to achieve the trade-offs between minimizing the sparse cost and the regrets for constraint violations. For large enough ρ value, e.g., $\rho \to \infty$, program* (3) *recovers to* SSCOP_N^1 (2). *In practice, parameter ρ must larger than the dual norm of the original Lagrange multiplier for the constraint* $\max_i \ h(x, q^i)$.

By virtue of Assumptions 1 and 2, the program (3) is tractable, and a general theory for scenario (linear) program with relaxation (but not involves sparse cost) has been proven in [4, Theroem 4], stated as follows

Lemma 1 (Violation [4]) *Under Assumptions 1 and 2, consider sparse optimization with relaxation problem* (3), *given a confidence $\beta \in (0,1)$, the violation of probability (i.e., risk) $V(x_N^*)$ is evaluated as follows*

$$\mathbb{P}^N\{\underline{\epsilon}(s_N^*) \le V(x_N^*) \le \overline{\epsilon}(s_N^*)\} \ge 1 - \beta \tag{4}$$

where $\underline{\epsilon}(\cdot) = \max\{0, 1 - \overline{t}(k)\}$, $\overline{\epsilon}(\cdot) = \max\{0, 1 - \underline{t}(k)\}$, s_N^* is the number of samples q^i for which $h(x, q^i) \ge 0$ at $x = x_N^*$, and for $k = 0, 1, \cdots, N-1$ the pair $\{\underline{t}(k), \overline{t}(k)\}$ is the solution of the polynomial equation in t variable

$$\mathfrak{B}_N(t;k) = \frac{\beta}{2N} \sum_{j=k}^{N-1} \mathfrak{B}_j(t;k) + \frac{\beta}{6N} \sum_{j=N+1}^{4N} \mathfrak{B}_j(t;k) \tag{5}$$

where $\mathfrak{B}_j(t;k) = \binom{j}{k} t^{j-k}$ is a binomial expansion. For $k = N$, the upper bound is set to $\overline{\epsilon}(k) = 1$ and the lower bound is derived by

$$1 = \frac{\beta}{6N} \sum_{j=N+1}^{4N} \mathfrak{B}_j(t;N). \tag{6}$$

Lemma 1 establishes an interval $[\underline{\epsilon}(s_N^*), \overline{\epsilon}(s_N^*)]$ for violation $V(x_N^*)$ that lies with a high confidence level $1-\beta$. The value of complexity s_N^* can be calculated according to the observed data samples $q^1, \cdots, q^N$, which is equivalent to the number of violated constraints plus at most n scenarios (i.e., the number of decision variables). Fig. 1 depicts the curves $\underline{\epsilon}(k)$ and $\overline{\epsilon}(k)$ for risk $V(x_N^*)$ with $N = 1000$, and $\beta = 10^{-5}, 10^{-7}, 10^{-9}$, respectively.

We next handle the risk-aware spare optimal control problem taken from sparse robust control [3], which is the main focus of attention in the present context.

3 Risk-Aware Sparse Optimal Control

In this section, we shift the principle from SSCOP_N^ρ to sparse optimal control problem, where the optimal control problem attempts to make the trade-offs between the sparse performance and violation constraints.

Consider the discrete-time uncertain linear time invariant (LTI) system over a finite time horizon

$$z_{l+1} = A(q)z_l + B(q)u_l, \quad l = 0, 1, \cdots, L-1, \tag{7}$$

where $z_l \in \mathbb{R}^n$ is the state with an initial condition z_0, $u_l \in \mathbb{R}^m$ is the control input, and the system matrices $A(q) \in \mathbb{R}^{n\times n}$, $B(q) \in \mathbb{R}^{n\times m}$ depend on the model parametric uncertainties $q \in \mathcal{Q} \subseteq \mathbb{R}^{n_q}$. Besides, assume that the pair $(A(q), B(q))$ is reachable.

For uncertain LTI dynamics (7), the control objective aims to seek a control sequence $\{u_l\}_{l=0}^{L-1}$ with input sparsity that drives the system behavior z_l from the initial state z_0 near to the terminal state z_L as well as restrain the uncertainty propagation q.

With the help of recursive iterations for the plant (7), the terminal state z_L can eventually be written as a scalar-valued uncertain function $h(u, q) : \mathbb{R}^{mL} \times \mathcal{Q} \to \mathbb{R}$ that measures a distance between the final state $z_L(u, q)$ and a prescribed target $\bar{z}$ (e.g., taking the origin $\bar{z} = 0$) with a radius $\gamma \geq 0$, defined by

$$\begin{aligned} h(u,q) &\doteq \|z_L(u,q) - \bar{z}\|_2 - \gamma \\ &= \|A(q)^L z_0 + \mathfrak{R}(q)u - \bar{z}\|_2 - \gamma \end{aligned} \tag{8}$$

with

$$\begin{aligned} \mathfrak{R}(q) &= \begin{bmatrix} A(q)^{L-1}B(q) & \ldots & A(q)B(q) & B(q) \end{bmatrix} \in \mathbb{R}^{n\times mL}, \\ u &= \begin{bmatrix} u_0^\top & u_1^\top & \ldots & u_{L-1}^\top \end{bmatrix}^\top \in \mathbb{R}^{mL}. \end{aligned}$$

Next, consider the risk-aware sparse optimal control problem that minimizes the sparse control inputs as well as the violated constraints subject to N random scenarios, built from program (3). In comparison with the existing work [3], we here make a trade-off between the sparse control cost and the feasibility by relaxing the constraint violations. Such problem appears when the designer needs to evaluate the control performance and risk level, thus we introduce the following risk-aware sparse optimal control problem.

Problem 1 (Risk-aware Sparse Control) *Given a discrete-time reachable uncertain LTI system* (7)*, and the parameters* $z_0, \bar{z}, L, \rho, \gamma$*, the risk-aware sparse optimal control is to achieve the trade-off between the sparse inputs and the risk assessment, which amounts to solving the following revised program* SSCOP_N^ρ*, minimizing*

$$\begin{aligned} \min_{u\in\mathbb{R}^{mL}, \xi^i \geq 0} \quad & \|u\|_1 + \rho \sum_{i=1}^{N} \xi^i \\ \text{s.t.} \quad & h(u, q^i) \leq \xi^i, \quad i = 1, \cdots, N. \end{aligned} \tag{9}$$

Then we formally state the following result, which is a restatement of the results of Lemma 1 in our context.

Theorem 1 (Violation of SSCOP_N^ρ) *With* $\underline{\epsilon}(\cdot)$ *and* $\overline{\epsilon}(\cdot)$ *in Lemma 1, given a confidence* $\beta \in (0, 1)$*, the risk assessment for violation* $V(u_N^*)$ *in risk-aware sparse optimal control problem* (9) *is as follows*

$$\mathbb{P}^N\{\underline{\epsilon}(s_N^*) \leq V(u_N^*) \leq \overline{\epsilon}(s_N^*)\} \geq 1 - \beta. \tag{10}$$

where s_N^* *is defined as the number of the violated constraints* $h(u_N^*, q^i) \geq 0$.

Theorem 1 offers the lower and upper bounds for risk-aware sparse optimal control on its violation $V(u_N^*)$, which holds with a high confidence level $1-\beta$. Besides, s_N^* indicates the violated constraints $h(u_N^*, q^i) > 0$, adding those active constraints $h(u_N^*, q^i) = 0$.

Remark 2 (Risk Estimation) *A common practice to estimate the risk of the sparse solution* $u_{N_\mathrm{T}}^*$ *based on the observations* q^i*, where* $i \in \mathcal{N}_\mathrm{T} = \{1, \cdots, N_\mathrm{T}\}$ *is the training samples, and* $i \in \mathcal{N}_\mathrm{V} = \{N_\mathrm{T}+1, \cdots, N_\mathrm{T}+N_\mathrm{V}\}$ *stands for the validation or out-of-sample samples, then the violation* $V(u_{N_\mathrm{T}}^*)$ *is estimated form the ratio*

$$\frac{\#\{i,\ h(u_{N_\mathrm{T}}^*, q^i) > 0,\ i = N_\mathrm{T}+1, \cdots, N_\mathrm{T}+N_\mathrm{V}\}}{N_\mathrm{V}}.$$

4 Numerical Simulations

In this section, we illustrate the numerical example to assess the risk for the proposed risk-aware sparse optimal control of Problem 1. Let us consider an uncertain mass-spring system, which can be modeled by a linear system (7) using ZoH sampling Ts = 0.05, with

$$A(q) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-(k_1+k_2)}{m_1} & 0 & \frac{k_2}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_2}{m_1} & 0 & \frac{-(k_2+k_3)}{m_2} & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ \frac{1}{m_1} & 0 \\ 0 & 0 \\ 0 & \frac{1}{m_2} \end{bmatrix}$$

where $m_1 = 1$, $m_2 = 2$ are the masses, and $k_1, k_2, k_3 \in \mathbb{R}$ are the flexibility of three springs, respectively. They can be collected by a vector $q = [k_1\ k_2\ k_3]^\top \in \mathcal{Q}$ to represent the uncertain parameters. Besides, the elements of the uncertain vector q are i.i.d. randomly sampled

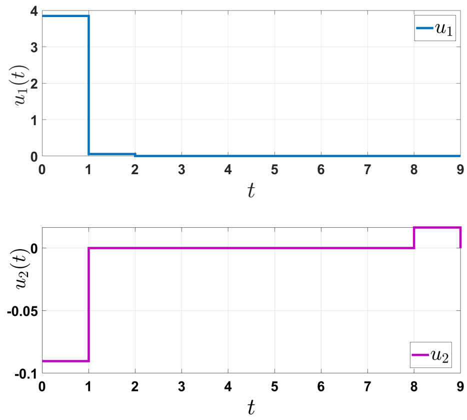

Fig. 2: Risk-aware sparse optimal control inputs.

from a uniform distribution over the interval $[0.1, 1]$. For this system, we compute the risk-aware sparse optimal control of Problem 1 over time horizon $L = 10$ with random samples $N = 1000$, and select the initial condition as $z_0 = [1\ 0\ 1\ 0]^\top$.

The objective is to seek a risk-aware control input u such that steers the terminal state $z_L(u, q)$ near to the target with a specified gap γ. We thus assign the target as the origin $\bar{z} = 0$ and set a small gap metric $\gamma = 0.5$ in (8). Note that this example is a convex optimization under randomized platform, and hence we applied the off-the-shelf package CVX in MATLAB to Problem 1.

Fig. 2 displays the risk-aware sparse optimal control with two inputs. It is clearly that the obtained optimal control signals (resp., the above and bottom) are truly sparse, and the computed optimal value was $\|u_N^*\|_1 + \frac{1}{5}\sum_i \xi_*^i = 247.083$. Fig. 3 depicts the risk-aware sparse optimal control under different penalty weights $\rho = 0.5, 1, 5$, respectively. From this figure, we can see that both of two control inputs are sparse when choose the appropriate value ρ. By tuning the parameter ρ, the user would make the desired trade-off between the sparse performance and the violated constraints.

Fig. 4 describes the sparse controls with additional scenarios validation $i = 1, \cdots, 2000$, it shows that the control signals enjoy sparsity and posterioi guarantees.

5 Conclusions

In this paper, we proposed the sparse optimization with relaxation and its application to risk-aware sparse optimal control for uncertain linear discrete-time system. We first utilized the scenario optimization to deal with the uncertainty propagation by random sampling. We then performed a relaxation for the constraints that makes a trade-off between the sparse cost and the violated constraints. Finally, we have simulated the numerical examples which demonstrate the effectiveness for the proposed risk-aware sparse optimal control.

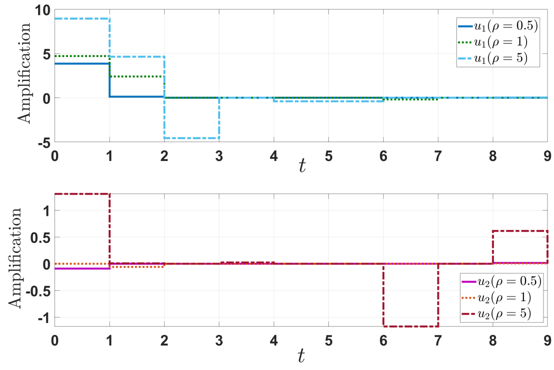

Fig. 3: Risk-aware sparse optimal control inputs with different penalty weights $\rho = 0.5, 1, 5$, respectively.

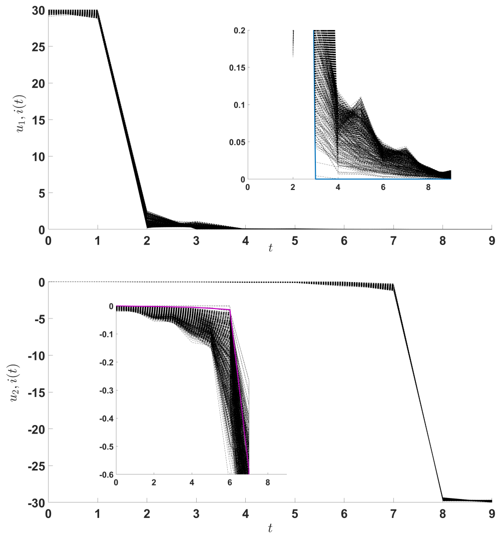

Fig. 4: Risk-aware sparse optimal control inputs with additional scenarios $N_V = 2000$.

References

[1] M. Nagahara, D. E. Quevedo, and D. Nešić, "Maximum hands-off control: a paradigm of control effort minimization," *IEEE Trans. Automa. Contr.*, vol. 61, no. 3, pp. 735–747, 2015.

[2] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Trans. Automa. Contr.*, vol. 51, no. 5, pp. 742–753, 2006.

[3] Z. Zhang and Y. Fujisaki, "Sparse robust control design via scenario program," *Proc. of the ISCIE Int. Sym. Stoch. Sys. Theory and its Appli*, vol. 2022, pp. 61–64, 2022.

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

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