Performance Bounds for the Scenario Approach and an Extension to a Class of Non-Convex Programs

Peyman Mohajerin Esfahani, Tobias Sutter, and John Lygeros

IEEE TAC, Vol. 60, No. 1, pp. 46–58, Jan. 2015

Paper Introduction

D1, Zhicheng Zhang

July 27, 2021

Goal

This paper focuses on

- Consider the scenario convex program (SCP) for (relaxed) robust convex program (RCP) and chance-constrained program (CCP).
- Consider the tail probability of the worst-case violation.
- Build a probabilistic bridge from the optimal value of SCP ($J^\star_N$) to the optimal value of RCP ($J^\star_{\mathrm{RCP}}$) and CCP ($J^\star_{\mathrm{CCP}}$).

Problem Statement

Technical Statements

- Let $\mathbb{X} \in \mathbb{R}^n$ be a compact convex set and $c \in \mathbb{R}^n$ a constant vector.
- Let $(\mathcal{D}, \mathcal{B}(\mathcal{D}), \mathbb{P})$ be probability space, where $\mathcal{D}$ is a metric space with the respective Borel $\sigma-$algebra $\mathcal{B}(\mathcal{D})$.
- A measurable function $f : \mathbb{X} \times \mathcal{D} \mapsto \mathbb{R}$, which is convex in the first argument and bounded in the second argument.

Let us consider the following optimization problems:

Robust Convex Program (RCP)

$$\begin{aligned} \min_{x \in \mathbb{X}} \quad & c^\top x \\ \text{s.t.} \quad & f(x, d) \leq 0 \quad \forall d \in \mathcal{D} \end{aligned}$$

where the optimal value of the RCP is denoted by $J^\star_{\mathrm{RCP}}$.

Relaxed (Perturbed) RCP_γ

Consider the Relaxed (Perturbed) RCP for $\gamma > 0$

$$\begin{aligned} \min_{x \in \mathbb{X}} \quad & c^\top x \\ \text{s.t.} \quad & f(x, d) \leq \gamma, \quad \forall d \in \mathcal{D} \end{aligned}$$

with the optimal value $J^\star_{\mathrm{RCP}_\gamma}$.

Chance-Constrained Program (CCP_ε)

$$\begin{aligned} \min_{x \in \mathbb{X}} \quad & c^\top x \\ \text{s.t.} \quad & \mathbb{P}\big[f(x, d) \leq 0\big] \geq 1 - \varepsilon \end{aligned}$$

with the optimal value $J^\star_{\mathrm{CCP}_\varepsilon}$.

- x is an ε-probability robust design, if it holds that $\mathbb{P}\big[f(x, d) > 0\big] \leq \varepsilon$.

Standard Scenario Convex Program [Campi & Garatti]

- Assume probability measure $\mathbb{P}$ on $\mathcal{D}$
- Sample $(d_i)_{i=1}^N$ independent and identically distributed (i.i.d.) from $\mathbb{P}$
- Formulate the scenario convex program

Scenario Convex Program (SCP)

$$\begin{aligned} \min_{x \in \mathbb{X}} \quad & c^\top x \\ \text{s.t.} \quad & f(x, d_i) \leq 0, \quad \forall i \in \{1, \cdots, N\} \end{aligned}$$

where the optimal solution and value of SCP are denoted by $x_N^\star$ and $J_N^\star$.

SP: Value Bound

Theorem 2.2 (CCP$_\varepsilon$ Feasibility)

Let $\beta \in [0,1]$ and $N \geq N(\varepsilon;\beta)$ where

$$N \geq N(\varepsilon;\beta) := \min\left\{ N \in \mathbb{N} \mid \sum_{i=0}^{n-1} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} \leq \beta \right\}$$

Then, the optimizer of SCP is a feasible solution of CCP$_\varepsilon$ with probability $\mathbb{P}^N$ (N-fold product probability measure) at least $1-\beta$.

- Alternatively, it states that $\mathbb{P}^N\left[x_N^\star \models \mathrm{CCP}_\varepsilon\right] \geq 1-\beta$, where $\models$ refers to the feasibility satisfaction, i.e., $x_N^\star \models \mathrm{CCP}_\varepsilon$ means that $x_N^\star$ is a feasible solution for the program CCP$_\varepsilon$.
- w.p. at least $1-\beta$, if SCP is feasible it holds that $J_N^\star \geq J^\star_{\mathrm{CCP}_\varepsilon}$
- It is sufficient to select a sample size $N \geq \frac{2}{\varepsilon}\left(n - 1 + \mathrm{In}\ \frac{1}{\beta}\right) \geq N(\varepsilon;\beta)$

Uniform Level-set Bound (ULB) [Kanamori & Takeda]

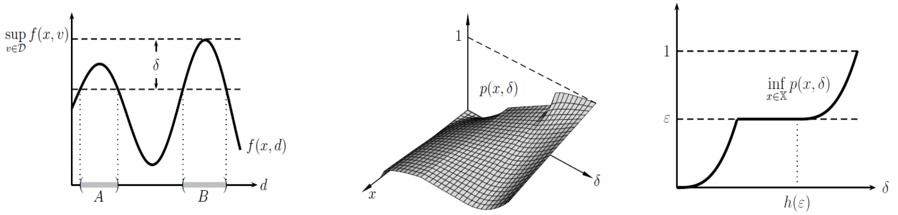
Definition 3.1

The *tail probability of the worst-case violation* $p : \mathbb{X} \times \mathbb{R}_+ \mapsto [0,1]$ is the function defined as

$$p(x,\delta) := \mathbb{P}\left[\sup_{v \in \mathcal{D}} f(x,v) - \delta < f(x,d)\right]$$

We call $h : [0,1] \mapsto \mathbb{R}_+$ a *uniform level-set bound* (ULB) of p if

$$h(\varepsilon) \geq \sup\left\{\delta \in \mathbb{R}_+ \mid \inf_{x \in \mathbb{X}} p(x,\delta) \leq \varepsilon\right\}, \quad \forall\, \varepsilon \in [0,1].$$



(a) $p(x,\delta) = \mathbb{P}[A \cup B]$

(b) Tail probability of the worst-case violation

(c) Uniform level set bound

Lemma 3.2

Let $h : [0,1] \mapsto \mathbb{R}_+$ be a ULB. Then

$$x \models \mathrm{CCP}_\varepsilon \Rightarrow x \models \mathrm{RCP}_{h(\varepsilon)}$$

that is, the feasible set of CCP_ε with constraint violation level ε is a subset of the feasible set of the relaxed program $\mathrm{RCP}_{h(\varepsilon)}$ with $\gamma := h(\varepsilon)$.

Proposition 3.8

Assume that the mapping $\mathcal{D} \ni d \mapsto f(x,d) \in \mathbb{R}$ is Lipschitz continuous with constant L_d uniformly in $x \in \mathbb{X}$. Suppose there exists a strictly increasing function $g : \mathbb{R}_+ \mapsto [0,1]$ such that

$$\mathbb{P}\Big[\mathbb{B}_r(d)\Big] \geq g(r), \quad \forall d \in \mathcal{D}$$

where $\mathbb{B}_r(d) \subset \mathcal{D}$ is an open ball centered at d with radius r. Then, $h(\varepsilon) := L_d g^{-1}(\varepsilon)$ is a ULB in the sense of Definition 3.1, where g^{-1} is the inverse function of g.

Bad news: Curse of dimensionality !

Key Assumption and Lemma

Assumption 3.3 (Slater Point)

There exists an $x_0 \in \mathbb{X}$ such that $\sup_{d \in \mathcal{D}} f(x,d) < 0$. Define the constant:

$$L_{SP} := \frac{\min_{x \in \mathbb{X}} c^\top x - c^\top x_0}{\sup_{d \in \mathcal{D}} f(x,d)}$$

Lemma 3.4

Consider the relaxed program RCP_γ and its optimal value $J^\star_{\mathrm{RCP}_\gamma}$. Under Assumption 3.3, the mapping $\mathbb{R}_+ \ni \gamma \mapsto J^\star_{\mathrm{RCP}_\gamma} \mathbb{R}$ is Lipschitz continuous with constant bounded by L_{SP} i.e., for all $\gamma_2 \geq \gamma_1 \geq 0$ we have

$$0 \leq J^\star_{\mathrm{RCP}_{\gamma_1}} - J^\star_{\mathrm{RCP}_{\gamma_2}} \leq L_{SP}(\gamma_2 - \gamma_1)$$

- *Proof:* Under the strong duality condition the mapping $\gamma \mapsto \mathrm{RCP}_\gamma$, which is Lipschitz continuous with the constant $\|\lambda^\star\|_1$ where $\lambda^\star$ is a dual optimizer of the RCP. In [33] it implies that $\|\lambda^\star\|_1 \leq L_{SP}$.
- Note that Lemma 3.4 also can be applied to the program SCP.

Main Results

Theorem 3.6 (RCP Confidence Interval)

Consider the programs RCP and SCP with the associated optimal values $J^\star_{\mathrm{RCP}}$ and $J^\star_N$, respectively. Suppose Assumption 3.3 holds and with constant L_{SP}. Given a ULB h and $\varepsilon, \beta \in [0,1]$ for all $N \geq N(\varepsilon, \beta)$, we have

$$\mathbb{P}^N\left[J^\star_{\mathrm{RCP}} - J^\star_N \in [0, I(\varepsilon)]\right] \geq 1 - \beta$$

where $I(\varepsilon) := \min\left\{L_{SP}h(\varepsilon), \max_{x\in\mathbb{X}} c^\top x - \min_{x\in\mathbb{X}} c^\top x\right\}$

- *Proof:* $\mathbb{P}^N\left[x^\star_N \models \mathrm{CCP}_\varepsilon\right] \geq 1 - \beta \Rightarrow^{Le3.2} \mathbb{P}^N\left[x^\star_N \models \mathrm{RCP}_{h(\varepsilon)}\right] \geq 1 - \beta$ $\Rightarrow \mathbb{P}^N\left[J^*_{\mathrm{RCP}_{h(\varepsilon)}} \leq J^*_N\right] \geq 1 - \beta \Rightarrow^{Le3.4} J^\star_{\mathrm{RCP}} \leq J^\star_{\mathrm{RCP}_{h(\varepsilon)}} + L_{SP}h(\varepsilon)$ Then, $\mathbb{P}^N\left[J^\star_{\mathrm{RCP}} \leq J^*_N + L_{SP}h(\varepsilon)\right] \geq 1 - \beta$. Meanwhile, SCP is a restricted version of RCP, then it has $J^\star_N \leq J^\star_{\mathrm{RCP}}$.
- Explicit bound: Under Pro 3.8, for any $N \geq N(g(\frac{\varepsilon}{L_{SP}L_d}), \beta)$, we have

$$\mathbb{P}^N\left[J^\star_{\mathrm{RCP}} - J^\star_N \in [0, \varepsilon]\right] \geq 1 - \beta$$

Main Results

Theorem 3.7 (CCP$_\varepsilon$ Confidence Interval)

Consider the programs CCP$_\varepsilon$ and SCP with the associated optimal values $J^\star_{\mathrm{CCP}_\varepsilon}$ and $J^\star_N$, respectively. Let h be a ULB and $\lambda^\star_N$ be the dual optimizer of SCP. Given $\beta \in [0,1]$ for all $N \geq N(\varepsilon, \beta)$, we have

$$\text{A Priori Assessment:} \quad \mathbb{P}^N\left[J^\star_{\mathrm{CCP}_\varepsilon} - J^\star_N \in [-I(\varepsilon), 0]\right] \geq 1 - \beta$$

$$\text{A Posteriori Assessment:} \quad \mathbb{P}^N\left[J^\star_{\mathrm{CCP}_\varepsilon} - J^\star_N \in [-I_N(\varepsilon), 0]\right] \geq 1 - \beta$$

where the *a priori* interval $I(\varepsilon)$ is defined in the former, and the *a posteriori* interval is $I_N(\varepsilon) := \min\left\{\|\lambda^\star_N\|_1 h(\varepsilon), \max_{x\in\mathbb{X}} c^\top x - \min_{x\in\mathbb{X}} c^\top x\right\}$.

- *Proof:* $\mathbb{P}^N\left[J^\star_{\mathrm{RCP}_{h(\varepsilon)}} \leq J^\star_{\mathrm{CCP}_\varepsilon} \leq J^\star_N\right] \geq 1 - \beta$. Lemma 3.4 ensures $J^\star_N \leq J^\star_{\mathrm{RCP}} \leq J^\star_{\mathrm{RCP}_{h(\varepsilon)}} + L_{SP} h(\varepsilon)$. The optimal value of scenario relaxed RCP$_{h(\varepsilon)}$ is denoted by $J^\star_{N,h(\varepsilon)}$. $J^\star_{N,h(\varepsilon)} \leq J^\star_{\mathrm{RCP}_{h(\varepsilon)}}$ w.p. 1. Approximate the L_{SP} as $\|\lambda^\star_N\|_1$, and apply Lemma 3.4 to SCP,

$$J^\star_N - \|\lambda^\star_N\|_1 h(\varepsilon) \leq J^\star_{N,h(\varepsilon)} \leq J^\star_{\mathrm{RCP}_{h(\varepsilon)}}$$

Feasibility of RCP via SCP

An Example to RCP & SCP

$$
\begin{aligned}
&\min_{x\in\mathbb{X}:=[-1,1]} && -x \\
\text{RCP:}\quad & \text{s.t.} && x-d\le 0,\quad \forall d\in\mathcal{D} \\
\text{SCP:}\quad & \text{s.t.} && x-d_i\le 0,\quad \forall i\in\{1,\cdots,N\}
\end{aligned}
$$

- The feasible set of RCP is $[-1,0]$ with the optimizer $x^\star = 0$.
- The optimizer of SCP is $x_N^\star = \min_{i\le N} d_i$.
- If probability $\mathbb{P}$ does not have atoms (point measure), we have $\mathbb{P}^N[\min_{i\le N} d_i > 0] = 1$. Thus, one can deduce that
$$\mathbb{P}^N[x_N^\star \models \text{RCP}] = 0,\quad \forall \mathbb{P}\in\mathcal{P},\quad \forall N\in\mathbb{N}$$
where $\mathcal{P}$ is the family of all non-atomic measure on $(\mathcal{D},\mathcal{B}(\mathcal{D}))$.
- If the set $\arg\max_{d\in\mathcal{D}} f(x,d)$ has measure zero for any $x\models\text{RCP}$, then SCP almost surely (a.s.) returns infeasible solutions to RCP, as worst-case scenarios are a.s. neglected.

Measurability of SCP Optimizer

Two-stage SCP

Let $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ is a strictly convex function. Then, consider the program

$$\begin{aligned} \min_{x \in \mathbb{X}} \quad & \phi(x) \\ \text{s.t.} \quad & f(x, d_i) \leq 0, \quad \forall i \in \{1, \cdots, N\} \\ & c^\top x \leq J_N^\star \end{aligned}$$

with the optimal solution (optimizer) $\tilde{x}_N^\star$.

- $f(x, d)$ with respect to the x is lower semi-continuous.
- Clearly, $\tilde{x}_N^\star$ is an optimizer of the program SCP.
- Proposition 3.10 (Measurability of the Optimizer)
- Proposition 3.11 (Measurability of the Feasible Set)
- Constraint function $f(x, d)$: measurability w.r.t. the uncertainty and lower semi-continuity w.r.t. the decision variables.

Example: Quadratic Constraints via Infinite Hyperplanes

Model Setup

- Take decision variable $x = [x_1, x_2]^\top \in \mathbb{X} : [0,1]^2 \subset \mathbb{R}^2$
- Define linear objective function $c := [-1, -1]^\top$
- Choose constraint function $f(x, d) := x_1 cos(d) + x_2 sin(d) - 1$
- $\mathbb{P}$ is the uniform probability measure on uncertainty $d \in \mathcal{D} := [0, 2\pi]$

Many infinite hyperplane constraints are rewritten as quadratic constraint

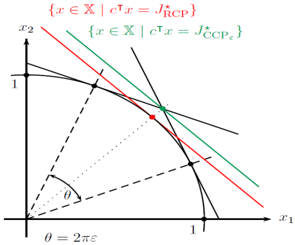
$$\max_{d \in [0,2\pi]} x_1 cos(d) + x_2 sin(d) - 1 \leq \gamma \iff x_1^2 + x_2^2 \leq (\gamma + 1)^2,\ \gamma \geq 0$$

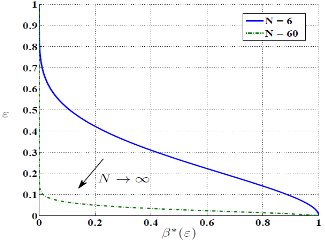
And the analytical solutions of optimal values for programs

$$J^\star_{\mathrm{RCP}_\gamma} = \max\{-\sqrt{2}(\gamma+1)\}, \quad J^\star_{\mathrm{CCP}_\varepsilon} = \max\{-\sqrt{2}/cos(\pi\varepsilon), -2\}$$

Fixed N scenarios for SCP with the optimal value $J^\star_N$, Given $\varepsilon, \beta \in [0,1]$,

$$\beta^*(\varepsilon) = \sum_{i=0}^{n-1} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i} = (1-\varepsilon)^N + N\varepsilon(1-\varepsilon)^{N-1},\ (i.e., n = 2)$$





- In this case $x_0 = [0, 0]^\top$ is Slater point, then $L_{SP} = \frac{-2-0}{-1} = 2$.
- The mapping $d \mapsto f(x, d)$ has Lipschitz constant $L_d = \sqrt{2}$ over $\mathbb{X}$.
- From Pro. 3.8, it derives that $g(r) = \frac{r}{\pi}$ and the ULB $h(\varepsilon) := \sqrt{2}\pi\varepsilon$.
- Then the confidence interval $I(\varepsilon) := \max\{2\sqrt{2}\pi\varepsilon, 2\}$, this arrives at $J^\star_{\mathrm{RCP}} - J^\star_N \in [0, I(\varepsilon)]$ (resp. $J^\star_{\mathrm{CCP}} - J^\star_N \in [-I(\varepsilon), 0]$) w.p. $1 - \beta^*(\varepsilon)$.

To validate this result, Solving program SCP for M different experiments:

- For each $k \in \{1, 2, \cdots, M\}$, drawing N scenarios $(d_i(k))_{k=1}^N \subset [0, 2\pi]$ w.r.t. uniform probability distribution $\mathbb{P}$, and then solve the SCP.
- Let $J^\star_N(k)$ be the optimal value of the kth experiment.

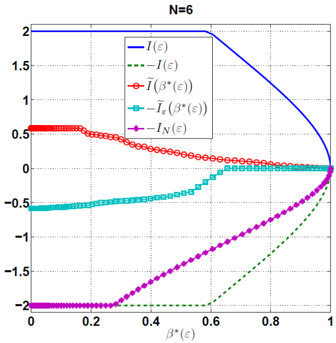
- Given $\beta \in [0,1]$, the empirical confidence interval of RCP can be represented by the minimal $\tilde{I}(\beta)$ so that the interval $[0, \tilde{I}(\beta)]$ contains $J_N^\star(m) - J_{\mathrm{RCP}}^\star$ for at least m experiments where $\frac{m}{M} \geq 1 - \beta$, i.e.,

$$\tilde{I}(\beta) := \min\Big\{\tilde{I} \in \mathbb{R}_+ \mid \exists A \subset \{1, \cdots, M\} : |A| \geq (1-\beta)M \\ \text{and } J_{\mathrm{RCP}}^\star - J_N^\star(k) \in [0, \tilde{I}] \quad \forall k \in A\Big\}$$

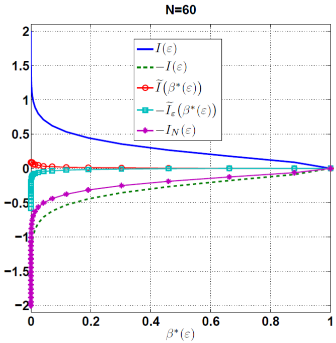
- Regarding CCP_ε, notice that the empirical confidence interval $\tilde{I}_\varepsilon(\beta)$ depends on both parameters ε and β since the analytical optimal values $J_{\mathrm{CCP}}^\star$ depends on ε as well. Then we can define

$$\tilde{I}_\varepsilon(\beta) := \min\Big\{\tilde{I} \in \mathbb{R}_+ \mid \exists A \subset \{1, \cdots, M\} : |A| \geq (1-\beta)M \\ \text{and } J_{\mathrm{CCP}_\varepsilon}^\star - J_N^\star(k) \in [-\tilde{I}, 0] \quad \forall k \in A\Big\}$$

- The sets $\tilde{I}(\beta)$ and $\tilde{I}_\varepsilon(\beta)$ are in close relation with sample quantiles.
- Set the number of experiments as $M = 2000$, scenarios $N = 6, 60$.
- The confidence interval $[0, I(\varepsilon)]$ (resp. $[-I(\varepsilon), 0]$) contains the empirical confidence interval $[0, \tilde{I}(\beta^*(\varepsilon))]$ (resp. $[-\tilde{I}_\varepsilon(\beta^*(\varepsilon)), 0]$)



(a) Simulations for $N = 6$



(b) Simulations for $N = 60$

- Choose one of the experiments and depict the corresponding *a posteriori* confidence interval $I_N(\varepsilon)$ versus $\beta^*(\varepsilon)$.
- Figs shows that a posteriori confidence interval $I_N(\varepsilon)$ (violet) proposes a tighter bound than the a priori confidence interval $I(\varepsilon)$ (green).
- We conjecture that in general the dual optimizer of SCP may happen to be a better approximation in comparison with the constant L_{SP}.

Conclusion

- Propose probabilistic performance bounds for both RCP and CCP_{ε} via SCP.
- Based on the tail probability of the worst-case constraint violation of the SCP solution and perturbation theory of convex optimization to give bounds.
- Given confidence bounds for the objective performance of RCPs and CCPs based on SCP.

Outlook in this work

- Study the derivation of ULBs, depending on the uncertainty set and the constraint functions.
- Investigate the relation between the constant L_{SP} and the dual optimizers of the program SCP.