ROBIST: Robust Optimization By Iterative Scenario Sampling and Statistical Testing

A Practical Scenario-Based Approach to Optimization Under Uncertainty

Justin Starreveld Guanyu Jin Dick den Hertog Roger Laeven

31-05-2023 (SIAM OP23)

Faculty of Economics and Business, University of Amsterdam

Outline

Introduction

Optimization under parametric uncertainty

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \ & g(\mathbf{x}) \\ \text{s.t. } & f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0, \end{aligned}$$

where:

- $\mathbf{x} \in \mathbb{R}^{n_\mathbf{x}}$ is the decision vector, defined on a closed convex feasible set $\mathcal{X}$
- $\tilde{\mathbf{z}} \in \mathbb{R}^{n_\mathbf{z}}$ is an uncertain parameter vector
- $g(\mathbf{x})$ and $f(\mathbf{x}, \tilde{\mathbf{z}})$ are scalar-valued functions that are convex in $\mathbf{x}$

Optimization under parametric uncertainty

$$\min_{\mathbf{x}\in\mathcal{X}} g(\mathbf{x})$$
$$\text{s.t. } f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0,$$

where:

- $\mathbf{x} \in \mathbb{R}^{n_\mathbf{x}}$ is the decision vector, defined on a closed convex feasible set $\mathcal{X}$
- $\tilde{\mathbf{z}} \in \mathbb{R}^{n_\mathbf{z}}$ is an uncertain parameter vector
- $g(\mathbf{x})$ and $f(\mathbf{x}, \tilde{\mathbf{z}})$ are scalar-valued functions that are convex in $\mathbf{x}$

$\Rightarrow$ How to deal with uncertain constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0$?

Three main approaches

How to deal with uncertain constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0$?

1. Stochastic Programming (Charnes & Cooper, 1959)

$$\mathbb{P}(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon$$

Three main approaches

How to deal with uncertain constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0$?

1. Stochastic Programming (Charnes & Cooper, 1959)

$$\mathbb{P}(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon$$

2. Robust Optimization (Ben-Tal & Nemirovski, 1998)

$$\sup_{\mathbf{z} \in \mathcal{U}} f(\mathbf{x}, \mathbf{z}) \leq 0$$

Three main approaches

How to deal with uncertain constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0$?

1. Stochastic Programming (Charnes & Cooper, 1959)

$$\mathbb{P}(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon$$

2. Robust Optimization (Ben-Tal & Nemirovski, 1998)

$$\sup_{\mathbf{z} \in \mathcal{U}} f(\mathbf{x}, \mathbf{z}) \leq 0$$

3. Scenario Optimization (Calafiore & Campi, 2005)

$$f(\mathbf{x}, \mathbf{z}^j) \leq 0 \quad \forall j \in \{1, \ldots, m\}$$

Practical limitations to stochastic programming approach

$$\mathbb{P}(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon$$

- Assumes $\mathbb{P}$ is known
- Even when $\mathbb{P}$ is known, still generally intractable (Shapiro & Nemirovski, 2005)

Practical limitations to robust optimization approach

$$\sup_{\mathbf{z} \in \mathcal{U}} f(\mathbf{x}, \mathbf{z}) \leq 0$$

- Computational tractability of reformulation is highly dependent on f and $\mathcal{U}$
  - May lead to a huge increase in the number of additional variables and constraints
  - If f is non-concave in $\mathbf{z}$, exact reformulations are known only for specific $\mathcal{U}$
- Can be difficult to determine appropriate shape and size of $\mathcal{U}$

Practical limitations to scenario optimization approach

$$f(\mathbf{x}, \mathbf{z}^j) \leq 0 \quad \forall j \in \{1, \ldots, m\}$$

- Required number of (randomly sampled) scenarios m quickly becomes prohibitively large for medium- and large-scale optimization problems

Three main approaches

How to deal with uncertain constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0$?

1. Stochastic Programming (Charnes & Cooper, 1959)

$$\mathbb{P}(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon$$

2. Robust Optimization (Ben-Tal & Nemirovski, 1998)

$$\sup_{\mathbf{z} \in \mathcal{U}} f(\mathbf{x}, \mathbf{z}) \leq 0$$

3. Scenario Optimization (Calafiore & Campi, 2005)

$$f(\mathbf{x}, \mathbf{z}^j) \leq 0 \quad \forall j \in \{1, \ldots, m\}$$

Three main approaches

How to deal with uncertain constraint $f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0$?

1. Stochastic Programming (Charnes & Cooper, 1959)

$$\mathbb{P}(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon$$

2. Robust Optimization (Ben-Tal & Nemirovski, 1998)

$$\sup_{\mathbf{z} \in \mathcal{U}} f(\mathbf{x}, \mathbf{z}) \leq 0$$

3. Scenario Optimization (Calafiore & Campi, 2005)

$$f(\mathbf{x}, \mathbf{z}^j) \leq 0 \quad \forall j \in \{1, \ldots, m\}$$

$\Rightarrow$ Can we develop a more practical approach?

Methodology

Illustrative Example

Example: toy problem from Yanıkoğlu & den Hertog (2013)

$$\max_{x_1, x_2 \leq 1} \quad x_1 + x_2$$
$$\text{s.t. } \tilde{z}_1 x_1 + \tilde{z}_2 x_2 \leq 1,$$

- Uncertain parameters $\tilde{z}_1$ and $\tilde{z}_2$ both uniformly distributed on $[-1, 1]$

Example: toy problem from Yanıkoğlu & den Hertog (2013)

$$(-)\min_{\underbrace{x_1, x_2 \leq 1}_{\mathcal{X}}} \underbrace{-(x_1 + x_2)}_{g(\mathbf{x})}$$

$$\text{s.t.} \quad \underbrace{\tilde{z}_1 x_1 + \tilde{z}_2 x_2 - 1}_{f(\mathbf{x}, \tilde{\mathbf{z}})} \leq 0,$$

- Uncertain parameters $\tilde{z}_1$ and $\tilde{z}_2$ both uniformly distributed on $[-1, 1]$

Example: toy problem from Yanıkoğlu & den Hertog (2013)

$$\max_{x_1, x_2 \leq 1} \; x_1 + x_2$$
$$\text{s.t. } \tilde{z}_1 x_1 + \tilde{z}_2 x_2 \leq 1,$$

- Uncertain parameters $\tilde{z}_1$ and $\tilde{z}_2$ both uniformly distributed on $[-1, 1]$
- Nominal case:
  - Expected parameter values $(\bar{z}_1, \bar{z}_2) = (0, 0)$
  - Nominal (optimal) solution $(\bar{x}_1, \bar{x}_2) = (1, 1)$

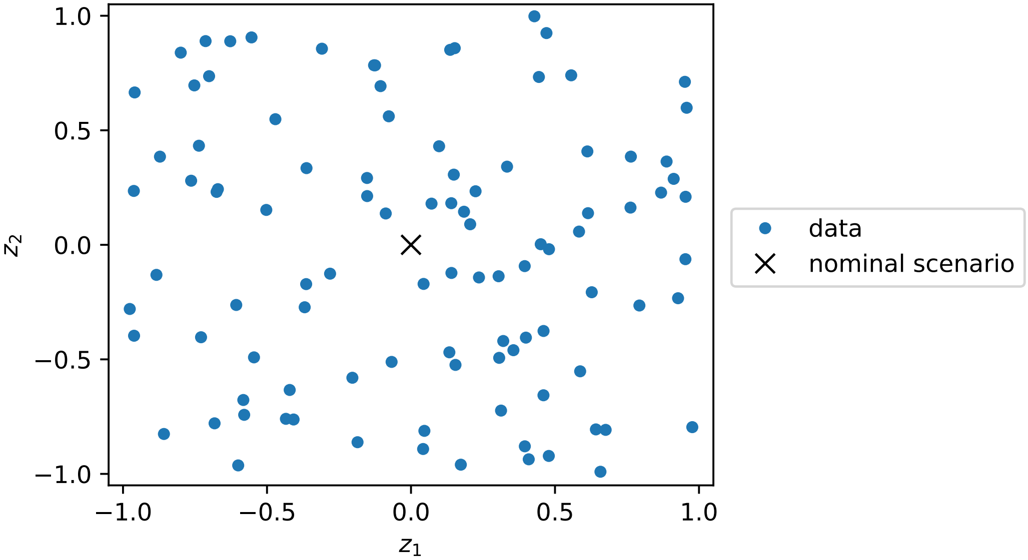
Example: toy problem from Yanıkoğlu & den Hertog (2013)

$$\max_{x_1, x_2 \leq 1} \; x_1 + x_2$$
$$\text{s.t. } \tilde{z}_1 x_1 + \tilde{z}_2 x_2 \leq 1,$$

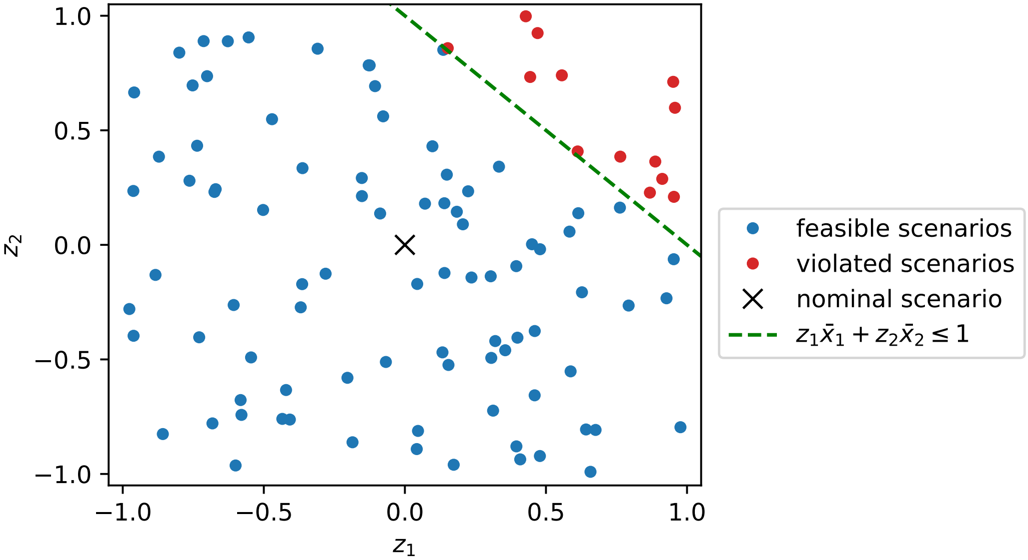
- Uncertain parameters $\tilde{z}_1$ and $\tilde{z}_2$ both uniformly distributed on $[-1, 1]$
- Nominal case:
  - Expected parameter values $(\bar{z}_1, \bar{z}_2) = (0, 0)$
  - Nominal (optimal) solution $(\bar{x}_1, \bar{x}_2) = (1, 1)$

$\Rightarrow$ How “robust” is this solution?

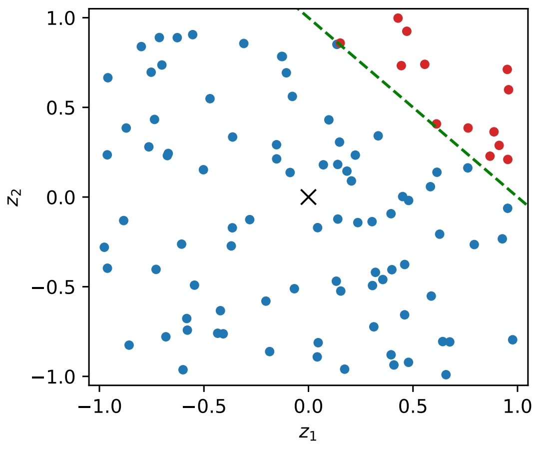
Imagine we have access to $N = 100$ scenarios/realizations of $(\tilde{z}_1, \tilde{z}_2)$



Analyze robustness of nominal solution $(\bar{x}_1, \bar{x}_2) = (1, 1)$ using our data

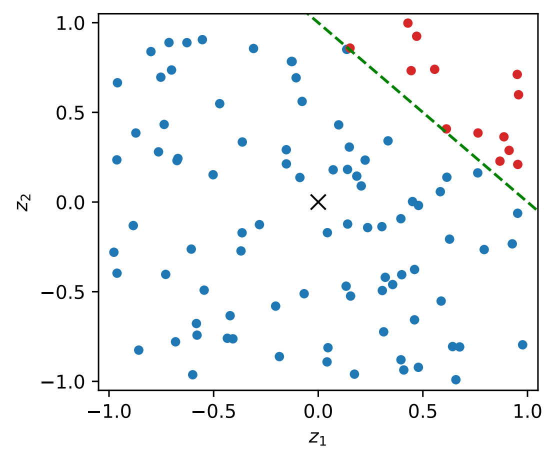


Use data to estimate the probability that x̄ is feasible



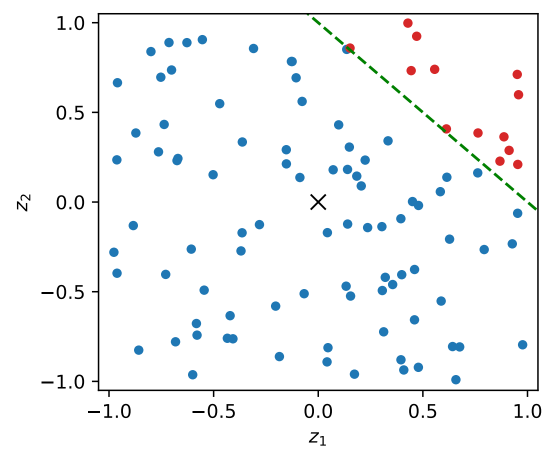
- We find that the uncertain constraint is violated for $\frac{13}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.87$

Use data to estimate the probability that $\bar{x}$ is feasible



- We find that the uncertain constraint is violated for $\frac{13}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.87$
- Derive statistical lower bound on estimate $(N = 100, \alpha = 0.01)$

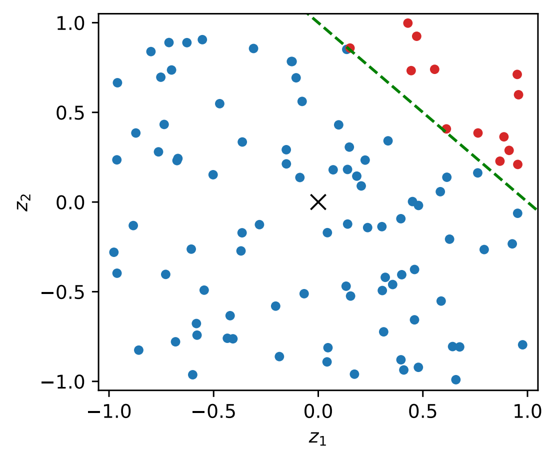
Use data to estimate the probability that x̄ is feasible



- We find that the uncertain constraint is violated for $\frac{13}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.87$
- Derive statistical lower bound on estimate $(N = 100, \alpha = 0.01)$

  $\Rightarrow$ Lower bound $= 0.78$

Use data to estimate the probability that $\bar{x}$ is feasible



- We find that the uncertain constraint is violated for $\frac{13}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.87$
- Derive statistical lower bound on estimate $(N = 100, \alpha = 0.01)$

  $\Rightarrow$ Lower bound $= 0.78$
- While optimizing for a single scenario $(\bar{z}_1, \bar{z}_2) = (0, 0)$, the resulting solution is likely to be feasible w.p. ≥ 0.78

Obtaining a more robust solution

- Imagine we are not content with the robustness of our nominal solution $(\bar{x}_1, \bar{x}_2)$
  - Feasible with probability ≥ 0.78 considered too risky

Obtaining a more robust solution

- Imagine we are not content with the robustness of our nominal solution $(\bar{x}_1, \bar{x}_2)$
  - Feasible with probability ≥ 0.78 considered too risky
- We can add scenarios/constraints to our optimization problem

$$
\begin{aligned}
\max_{x_1, x_2 \leq 1} \quad & x_1 + x_2 \\
\text{s.t.} \quad & \bar{z}_1 x_1 + \bar{z}_2 x_2 \leq 1, && \text{(nominal scenario)} \\
& z_1^i x_1 + z_2^i x_2 \leq 1, && \text{(scenario } i\text{)}
\end{aligned}
$$

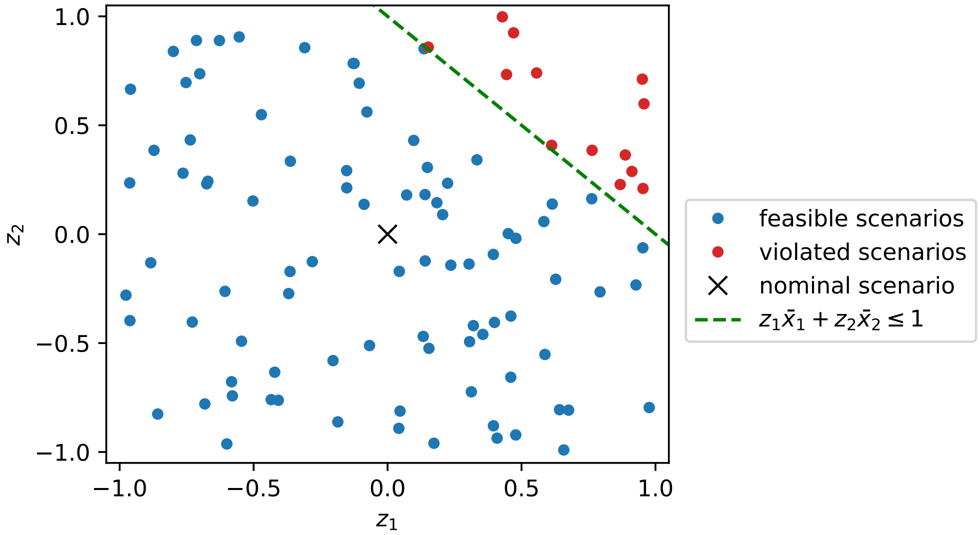
Obtaining a more robust solution

- Imagine we are not content with the robustness of our nominal solution $(\bar{x}_1, \bar{x}_2)$
  - Feasible with probability ≥ 0.78 considered too risky
- We can add scenarios/constraints to our optimization problem

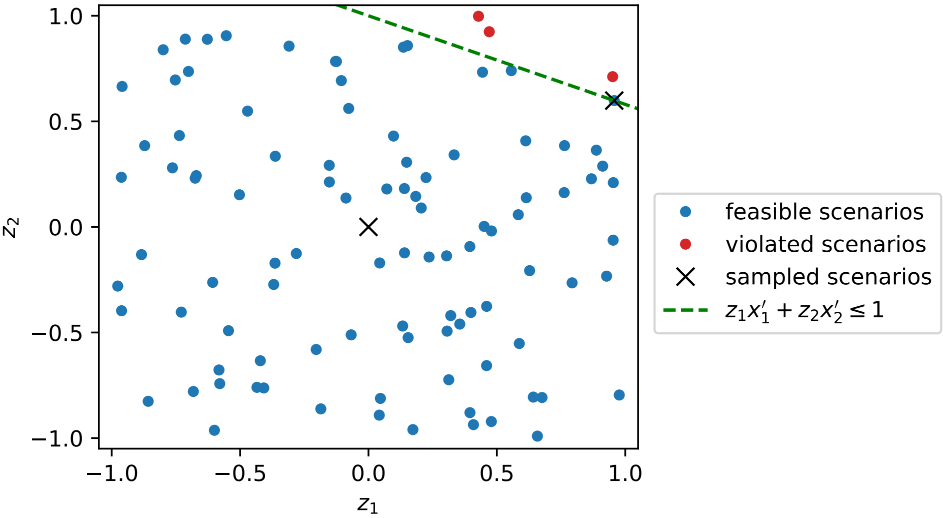
$$\begin{aligned} \max_{x_1, x_2 \leq 1} \quad & x_1 + x_2 \\ \text{s.t.} \quad & \bar{z}_1 x_1 + \bar{z}_2 x_2 \leq 1, && \text{(nominal scenario)} \\ & z_1^i x_1 + z_2^i x_2 \leq 1, && \text{(scenario } i\text{)} \end{aligned}$$

- Restricts the feasible region, and may lower the optimal objective value

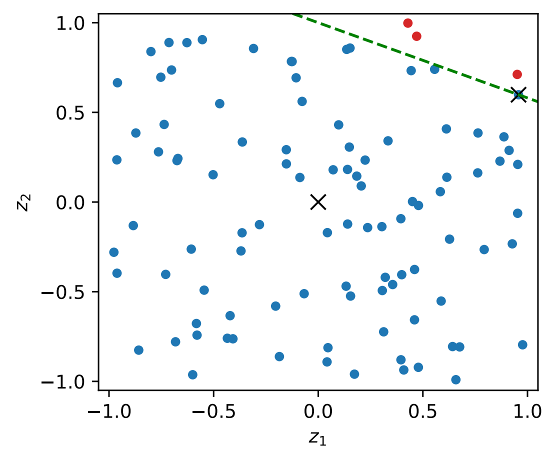
Which scenario(s) should be added?



Pick scenario $\hat{\mathbf{z}}^{11} = (0.96, 0.60)$ and resolve problem with added constraint $\Rightarrow \mathbf{x}' = (0.42, 1)$

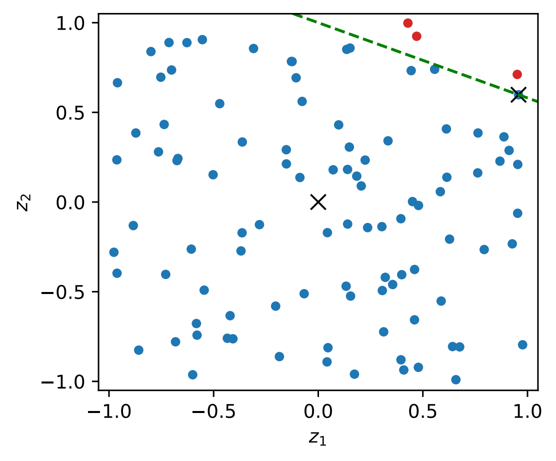


Analyze robustness of new solution $\mathbf{x}'$



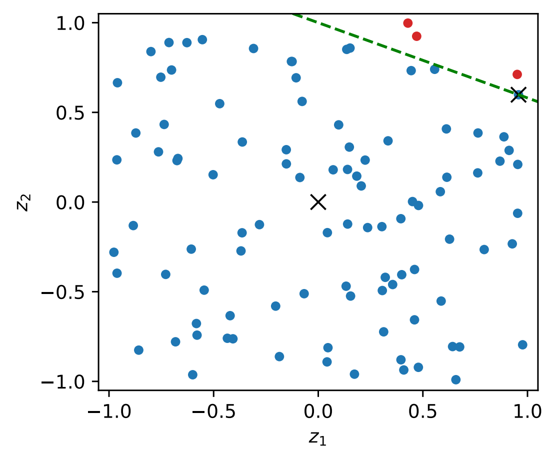
- We find that the uncertain constraint is violated for $\frac{3}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.97$

Analyze robustness of new solution $\mathbf{x}'$



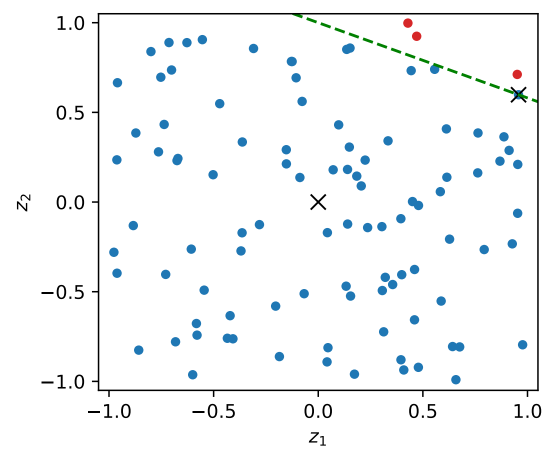
- We find that the uncertain constraint is violated for $\frac{3}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.97$
- Derive statistical lower bound on estimate $(N = 100, \alpha = 0.01)$

Analyze robustness of new solution $\mathbf{x}'$



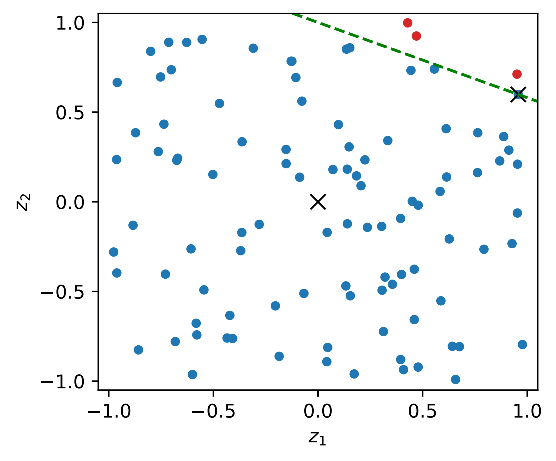
- We find that the uncertain constraint is violated for $\frac{3}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.97$
- Derive statistical lower bound on estimate $(N = 100, \alpha = 0.01)$
  - $\Rightarrow$ Lower bound $= 0.93$

Analyze robustness of new solution x′



- We find that the uncertain constraint is violated for $\frac{3}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.97$
- Derive statistical lower bound on estimate $(N = 100, \alpha = 0.01)$
  - $\Rightarrow$ Lower bound $= 0.93$
- Probability guarantee: $0.78 \rightarrow 0.93$

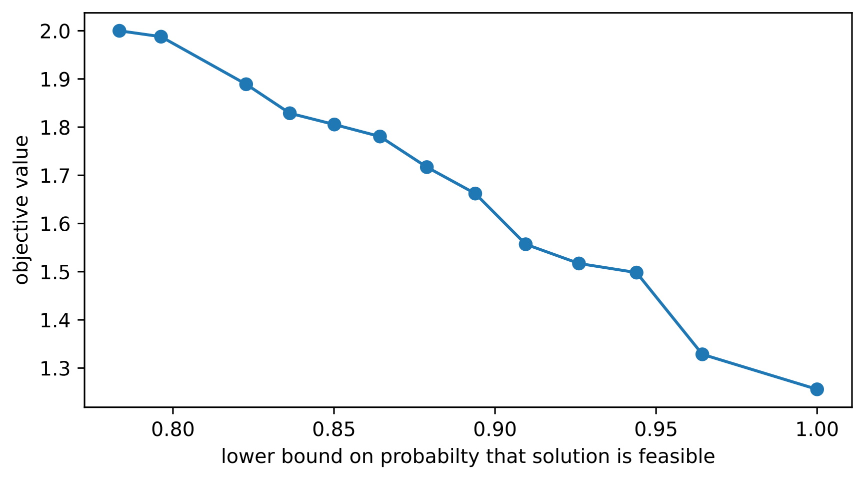
Analyze robustness of new solution $\mathbf{x}'$



- We find that the uncertain constraint is violated for $\frac{3}{100}$ of our scenarios
  - Empirical estimate of $\mathbb{P}^*(\tilde{z}_1\bar{x}_1 + \tilde{z}_2\bar{x}_2 \leq 1) = 0.97$
- Derive statistical lower bound on estimate $(N = 100, \alpha = 0.01)$
  - $\Rightarrow$ Lower bound $= 0.93$
- Probability guarantee: $0.78 \to 0.93$
- Objective value: $2.00 \to 1.42$

Trade-off between objective value and robustness

- Can construct trade-off curve from obtained solutions
  - Offers insight into "price of robustness" (Bertsimas & Sim, 2004)



Deriving probability guarantee

Problem setup

Uncertain convex program (UCP):

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{X}} \; & g(\mathbf{x}) \\ \text{s.t. } & f(\mathbf{x},\tilde{\mathbf{z}}) \leq 0, \end{aligned} \tag{UCP}$$

- We are interested in finding "robust" solutions to (UCP), i.e. solutions which are likely to be feasible despite the uncertainty

Problem setup

Uncertain convex program (UCP):

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{X}} \ & g(\mathbf{x}) \\ \text{s.t. } & f(\mathbf{x}, \tilde{\mathbf{z}}) \le 0, \end{aligned} \tag{UCP}$$

- We are interested in finding "robust" solutions to (UCP), i.e. solutions which are likely to be feasible despite the uncertainty
- Assume that $\tilde{\mathbf{z}}$ is a random variable with probability distribution $\mathbb{P}^*$

Problem setup

Uncertain convex program (UCP):

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \; & g(\mathbf{x}) \\ \text{s.t. } & f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0, \end{aligned} \tag{UCP}$$

- We are interested in finding "robust" solutions to (UCP), i.e. solutions which are likely to be feasible despite the uncertainty
- Assume that $\tilde{\mathbf{z}}$ is a random variable with probability distribution $\mathbb{P}^*$
- Given an tolerable probability of constraint violation ϵ, we would like the following "probability guarantee" to hold:

$$\mathbb{P}^*(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon.$$

Data + statistical testing to provide probability guarantee

- Assume we do not know $\mathbb{P}^*$, but have access to N i.i.d. data points $\mathcal{D} = \left\{\mathbf{z}^1, \ldots, \mathbf{z}^N\right\}$ that originate from $\mathbb{P}^*$

Data + statistical testing to provide probability guarantee

- Assume we do not know $\mathbb{P}^*$, but have access to N i.i.d. data points $\mathcal{D} = \left\{\mathbf{z}^1, \dots, \mathbf{z}^N\right\}$ that originate from $\mathbb{P}^*$
- Given such data and some fixed solution $\bar{\mathbf{x}}$, we can use statistical testing to state with confidence $\geq 1 - \alpha$ that:

$$\mathbb{P}^*(f(\bar{\mathbf{x}}, \tilde{\mathbf{z}}) \leq 0) \geq \bar{\gamma}$$

Data + statistical testing to provide probability guarantee

- Assume we do not know $\mathbb{P}^*$, but have access to N i.i.d. data points $\mathcal{D} = \left\{\mathbf{z}^1, \dots, \mathbf{z}^N\right\}$ that originate from $\mathbb{P}^*$
- Given such data and some fixed solution $\bar{\mathbf{x}}$, we can use statistical testing to state with confidence $\geq 1 - \alpha$ that:

$$\mathbb{P}^*(f(\bar{\mathbf{x}}, \tilde{\mathbf{z}}) \leq 0) \geq \bar{\gamma}$$

where:

- α is statistical probability of making a type I error

Data + statistical testing to provide probability guarantee

- Assume we do not know $\mathbb{P}^*$, but have access to N i.i.d. data points $\mathcal{D} = \{\mathbf{z}^1, \ldots, \mathbf{z}^N\}$ that originate from $\mathbb{P}^*$
- Given such data and some fixed solution $\bar{\mathbf{x}}$, we can use statistical testing to state with confidence $\geq 1 - \alpha$ that:

$$\mathbb{P}^*(f(\bar{\mathbf{x}}, \tilde{\mathbf{z}}) \leq 0) \geq \bar{\gamma}$$

where:

- α is statistical probability of making a type I error
- $\bar{\gamma}$ is the feasibility "certificate"

Our novel method for deriving feasibility "certificate" $\bar{\gamma}$

- Given some fixed solution $\bar{\mathbf{x}}$ and data set $\mathcal{D} = \left\{\mathbf{z}^1, \dots, \mathbf{z}^N\right\}$

Our novel method for deriving feasibility "certificate" $\bar{\gamma}$

- Given some fixed solution $\bar{\mathbf{x}}$ and data set $\mathcal{D} = \{\mathbf{z}^1, \ldots, \mathbf{z}^N\}$
- For each scenario $\mathbf{z}^i \in \mathcal{D}_N$, we compute whether $f(\bar{\mathbf{x}}, \mathbf{z}^i) \leq 0$ is satisfied

Our novel method for deriving feasibility "certificate" $\bar{\gamma}$

- Given some fixed solution $\bar{\mathbf{x}}$ and data set $\mathcal{D} = \{\mathbf{z}^1, \ldots, \mathbf{z}^N\}$
- For each scenario $\mathbf{z}^i \in \mathcal{D}_N$, we compute whether $f(\bar{\mathbf{x}}, \mathbf{z}^i) \leq 0$ is satisfied
    - Gives an empirical estimate $p_1 := \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{[f(\bar{\mathbf{x}}, \mathbf{z}^i) \leq 0]}$

Our novel method for deriving feasibility "certificate" $\bar{\gamma}$

- Given some fixed solution $\bar{\mathbf{x}}$ and data set $\mathcal{D} = \{\mathbf{z}^1, \ldots, \mathbf{z}^N\}$
- For each scenario $\mathbf{z}^i \in \mathcal{D}_N$, we compute whether $f(\bar{\mathbf{x}}, \mathbf{z}^i) \leq 0$ is satisfied
  - Gives an empirical estimate $p_1 := \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{[f(\bar{\mathbf{x}}, \mathbf{z}^i) \leq 0]}$
- Construct a $(1-\alpha)$-statistical confidence set $\mathcal{Q}_\phi$ around empirical estimate p_1 using the modified chi-squared distance

$$\mathcal{Q}_\phi(p_1, N, \alpha) := \left\{ q_1 \in \mathbb{R} \;\middle|\; q_1 \geq 0,\ \frac{(q_1 - p_1)^2}{p_1} + \frac{(q_1 - p_1)^2}{1 - p_1} \leq \frac{\chi^2_{1,1-\alpha}}{N} \right\}. \quad (1)$$

Our novel method for deriving feasibility "certificate" $\bar{\gamma}$

- Given some fixed solution $\bar{\mathbf{x}}$ and data set $\mathcal{D} = \left\{\mathbf{z}^1, \ldots, \mathbf{z}^N\right\}$
- For each scenario $\mathbf{z}^i \in \mathcal{D}_N$, we compute whether $f(\bar{\mathbf{x}}, \mathbf{z}^i) \leq 0$ is satisfied
  - Gives an empirical estimate $p_1 := \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{[f(\bar{\mathbf{x}}, \mathbf{z}^i) \leq 0]}$
- Construct a $(1-\alpha)$-statistical confidence set $\mathcal{Q}_\phi$ around empirical estimate p_1 using the modified chi-squared distance

$$\mathcal{Q}_\phi(p_1, N, \alpha) := \left\{ q_1 \in \mathbb{R} \;\middle|\; q_1 \geq 0,\; \frac{(q_1 - p_1)^2}{p_1} + \frac{(q_1 - p_1)^2}{1 - p_1} \leq \frac{\chi^2_{1,1-\alpha}}{N} \right\}. \quad (1)$$

- Determine certificate $\bar{\gamma}$ by computing $\bar{\gamma} := \min_{q_1 \in \mathcal{Q}_\phi} q_1$

Key elements in our novel method

- $\mathbb{P}^*(\tilde{\mathbf{z}} \in \mathcal{U})$ is an underestimation of $\mathbb{P}^*(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0)$
- Because we consider $\bar{\mathbf{x}}$ fixed and provide an "a posteriori" probability guarantee, the statistical test is on the *univariate* random variable $f(\bar{\mathbf{x}}, \tilde{\mathbf{z}})$
  - Two classes:
    1. $f(\bar{\mathbf{x}}, \tilde{\mathbf{z}}) \leq 0$
    2. $f(\bar{\mathbf{x}}, \tilde{\mathbf{z}}) > 0$

    $\Rightarrow$ Only 1 degree of freedom in statistical test

$\Rightarrow$ Sharp probability guarantees

$\Rightarrow$ Independent of dimensions of $\mathbf{x}$ and $\mathbf{z}$

Broad overview of our solution approach (ROBIST)

- Access to some sufficiently large data set $\mathcal{D}$
  - Split into training data $\mathcal{D}_{N_1}^{\text{train}}$ and testing data $\mathcal{D}_{N_2}^{\text{test}}$

Broad overview of our solution approach (ROBIST)

- Access to some sufficiently large data set $\mathcal{D}$
  - Split into training data $\mathcal{D}_{N_1}^{\text{train}}$ and testing data $\mathcal{D}_{N_2}^{\text{test}}$

1. Generation procedure
   - Sample a (small) subset of scenarios $\mathcal{S}$ from $\mathcal{D}_{N_1}^{\text{train}}$ to generate solutions

$$\mathbf{x} := \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} \{g(\mathbf{x}) \mid f(\mathbf{x}, \mathbf{z}) \leq 0, \ \forall \mathbf{z} \in \mathcal{S}\} \qquad (\text{SCP}_{\mathcal{S}})$$

Broad overview of our solution approach (ROBIST)

- Access to some sufficiently large data set $\mathcal{D}$
  - Split into training data $\mathcal{D}_{N_1}^{\text{train}}$ and testing data $\mathcal{D}_{N_2}^{\text{test}}$

1. Generation procedure
   - Sample a (small) subset of scenarios $\mathcal{S}$ from $\mathcal{D}_{N_1}^{\text{train}}$ to generate solutions

$$\mathbf{x} := \operatorname{argmin}_{\mathbf{x} \in \mathscr{X}} \{g(\mathbf{x}) \mid f(\mathbf{x}, \mathbf{z}) \leq 0, \ \forall \mathbf{z} \in \mathcal{S}\} \qquad (\text{SCP}_{\mathcal{S}})$$

2. Evaluation procedure
   - Use $\mathcal{D}_{N_2}^{\text{test}}$ to analyze the robustness of the generated solutions

Broad overview of our solution approach (ROBIST)

- Access to some sufficiently large data set $\mathcal{D}$
  - Split into training data $\mathcal{D}_{N_1}^{\text{train}}$ and testing data $\mathcal{D}_{N_2}^{\text{test}}$

1. Generation procedure
   - Sample a (small) subset of scenarios $\mathcal{S}$ from $\mathcal{D}_{N_1}^{\text{train}}$ to generate solutions

$$\mathbf{x} := \operatorname{argmin}_{\mathbf{x} \in \mathscr{X}} \{g(\mathbf{x}) \mid f(\mathbf{x}, \mathbf{z}) \leq 0, \ \forall \mathbf{z} \in \mathcal{S}\} \qquad (\text{SCP}_{\mathcal{S}})$$

2. Evaluation procedure
   - Use $\mathcal{D}_{N_2}^{\text{test}}$ to analyze the robustness of the generated solutions

$\Rightarrow$ Can we combine these two procedures in an iterative heuristic search algorithm?

The algorithm

Input:

- Two data sets $\mathcal{D}_{N_1}^{\text{train}}$ and $\mathcal{D}_{N_2}^{\text{test}}$

The algorithm

Input:

- Two data sets $\mathcal{D}_{N_1}^{\text{train}}$ and $\mathcal{D}_{N_2}^{\text{test}}$
  - * Data sets are independent and observations in $\mathcal{D}_{N_2}^{\text{test}}$ are i.i.d.

The algorithm

Input:

- Two data sets $\mathcal{D}_{N_1}^{\text{train}}$ and $\mathcal{D}_{N_2}^{\text{test}}$
  - * Data sets are independent and observations in $\mathcal{D}_{N_2}^{\text{test}}$ are i.i.d.
- Acceptable probability of constraint violation ϵ

The algorithm

Input:

- Two data sets $\mathcal{D}_{N_1}^{\text{train}}$ and $\mathcal{D}_{N_2}^{\text{test}}$
  - * Data sets are independent and observations in $\mathcal{D}_{N_2}^{\text{test}}$ are i.i.d.
- Acceptable probability of constraint violation ϵ
- Statistical confidence level α

The algorithm

Input:

- Two data sets $\mathcal{D}_{N_1}^{\text{train}}$ and $\mathcal{D}_{N_2}^{\text{test}}$
  - * Data sets are independent and observations in $\mathcal{D}_{N_2}^{\text{test}}$ are i.i.d.
- Acceptable probability of constraint violation ϵ
- Statistical confidence level α
- Stopping criteria (time limit and/or maximum number of iterations)

The algorithm

Input:

- Two data sets $\mathcal{D}_{N_1}^{\text{train}}$ and $\mathcal{D}_{N_2}^{\text{test}}$
  - * Data sets are independent and observations in $\mathcal{D}_{N_2}^{\text{test}}$ are i.i.d.
- Acceptable probability of constraint violation ϵ
- Statistical confidence level α
- Stopping criteria (time limit and/or maximum number of iterations)

Actions (at each iteration i):

- Solve ($\text{SCP}_{\mathcal{S}_i}$) to obtain $\mathbf{x}_i$, where $\mathcal{S}_i \subseteq \mathcal{D}_{N_1}^{\text{train}}$
- Evaluate robustness of $\mathbf{x}_i$ using all scenarios in $\mathcal{D}_{N_1}^{\text{train}}$
- If insufficiently robust, add a scenario to $\mathcal{S}_i$, else, remove a scenario from $\mathcal{S}_i$

The algorithm

Input:

- Two data sets $\mathcal{D}_{N_1}^{\text{train}}$ and $\mathcal{D}_{N_2}^{\text{test}}$
  - * Data sets are independent and observations in $\mathcal{D}_{N_2}^{\text{test}}$ are i.i.d.
- Acceptable probability of constraint violation ϵ
- Statistical confidence level α
- Stopping criteria (time limit and/or maximum number of iterations)

Actions (at each iteration i):

- Solve ($\text{SCP}_{\mathcal{S}_i}$) to obtain $\mathbf{x}_i$, where $\mathcal{S}_i \subseteq \mathcal{D}_{N_1}^{\text{train}}$
- Evaluate robustness of $\mathbf{x}_i$ using all scenarios in $\mathcal{D}_{N_1}^{\text{train}}$
- If insufficiently robust, add a scenario to $\mathcal{S}_i$, else, remove a scenario from $\mathcal{S}_i$

Output:

- Use $\mathcal{D}_{N_2}^{\text{test}}$ to determine best found solution $\mathbf{x}_{i^*}$

Numerical Experiments

Applications considered in the paper

1. Toy problem
   - Comparison with Calafiore & Campi (2005) and Yanıkoğlu & den Hertog (2013)
2. Portfolio management problem
   - Comparison with robust optimization approach, the data-driven uncertainty sets presented in Bertsimas et al. (2018)
3. Weighted distribution problem
   - Comparison with scenario optimization approach, the methods of Calafiore & Campi (2005); Carè et al. (2014); Calafiore (2016) and Garatti et al. (2022)
4. Two-stage adaptive lot-sizing problem
   - Comparison with Vayanos et al. (2012)

Application I: Toy Problem

Comparison with Calafiore & Campi (2005) and Yanıkoğlu & den Hertog (2013)

Method proposed in Calafiore & Campi (2005)

Solve a sampled convex program (SCP), where each scenario $\mathbf{z}^j$ is randomly sampled:

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \; & g(\mathbf{x}) \\ \text{s.t. } & f(\mathbf{x}, \mathbf{z}^j) \le 0 \quad \forall j \in \{1, \ldots, m\}. \end{aligned} \tag{SCP}$$

Method proposed in Calafiore & Campi (2005)

Solve a sampled convex program (SCP), where each scenario $\mathbf{z}^j$ is randomly sampled:

$$\begin{aligned} \min_{\mathbf{x} \in \mathscr{X}} \ & g(\mathbf{x}) \\ \text{s.t. } & f(\mathbf{x}, \mathbf{z}^j) \leq 0 \quad \forall j \in \{1, \ldots, m\}. \end{aligned} \tag{SCP}$$

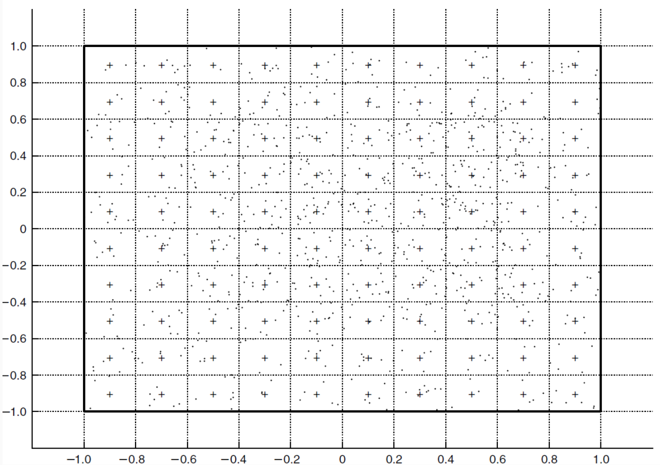
To ensure that the resulting solution satisfies $\mathbb{P}^*(f(\mathbf{x}, \tilde{\mathbf{z}}) \leq 0) \geq 1 - \epsilon$, with confidence $\geq 1 - \alpha$, the following must hold:

$$m \geq \frac{\dim(\mathbf{x})}{\epsilon\alpha} - 1,$$

This result was later tightened in Campi & Garatti (2008)

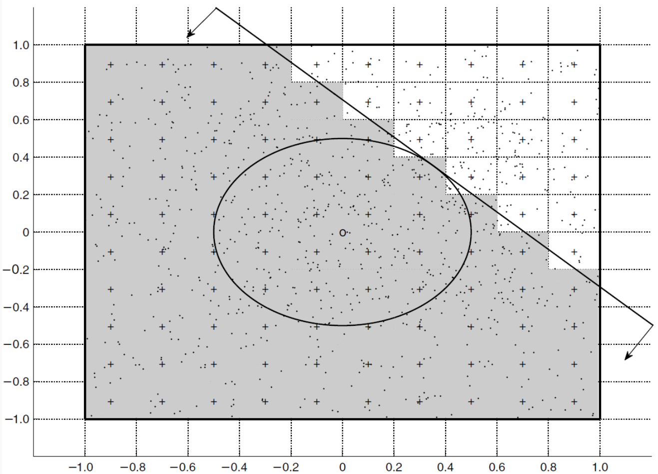
Method proposed in Yanıkoğlu & den Hertog (2013)

1) Divide uncertainty space into cells



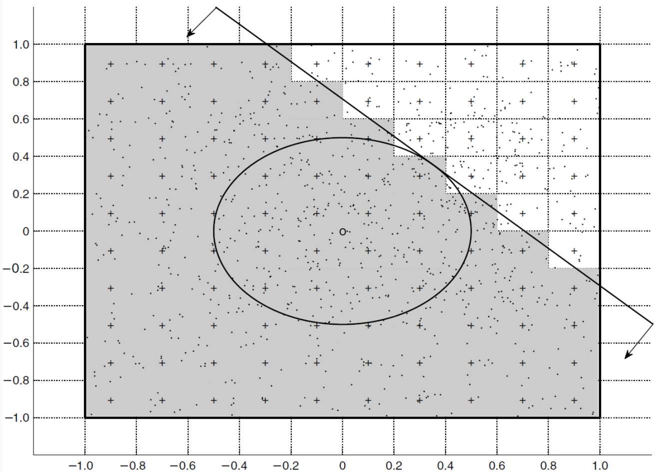
Method proposed in Yanıkoğlu & den Hertog (2013)

2) Solve robust counterpart with (ellipsoidal) uncertainty set



Method proposed in Yanıkoğlu & den Hertog (2013)

3) Compute probability guarantee with modified χ^2 test



Toy Problem

Same toy problem as earlier, but now in k dimensions:

$$\max_{\mathbf{x} \leq 1} \; \mathbf{e}^\intercal \mathbf{x} \tag{2}$$

$$\tilde{\mathbf{z}}^\intercal \mathbf{x} \leq 1, \tag{3}$$

where:

- $\mathbf{x} \in \mathbb{R}^k$
- $\tilde{\mathbf{z}} \in [-1, 1]^k$

Settings for numerical experiment

- Desired probability feasible $1 - \epsilon = 0.95$
- Statistical confidence level $1 - \alpha = 0.99$
- ROBIST settings:
  - $N_1 = N_2 = 500$
  - $i_{max} = 500$

Results: data and computation time

	N			Computation time (s)		
k	C&C	Y&dH	ROBIST	C&C	Y&dH	ROBIST
2	90	1,000	1,000	0.1	6	5
3	130	10,000	1,000	0.2	6	5
4	165	100,000	1,000	0.2	27	7
5	198	1,000,000	1,000	0.2	200	7
⋮						
10	344	10^{11}	1,000	0.4	-	12

Results: quality of solutions

	Objective			Probability guarantee		
k	C&C	Y&dH	ROBIST	C&C	Y&dH	ROBIST
2	1.19	1.20	1.32	0.950	0.969	0.951
3	1.39	1.42	1.63	0.950	0.958	0.951
4	1.57	1.67	1.83	0.950	0.952	0.951
5	1.76	1.85	2.06	0.950	0.951	0.952
⋮						
10	2.48	-	2.84	0.950	-	0.952

Results: quality of solutions

	Objective			Probability guarantee			Out-of-sample probability		
k	C&C	Y&dH	ROBIST	C&C	Y&dH	ROBIST	C&C	Y&dH	ROBIST
2	1.19	1.20	1.32	0.950	0.969	0.951	0.981	0.985	0.968
3	1.39	1.42	1.63	0.950	0.958	0.951	0.977	0.985	0.964
4	1.57	1.67	1.83	0.950	0.952	0.951	0.975	0.983	0.966
5	1.76	1.85	2.06	0.950	0.951	0.952	0.973	0.983	0.961
⋮									
10	2.48	-	2.84	0.950	-	0.952	0.972	-	0.959

Conclusion

Main takeaways

- While optimization under uncertainty can be difficult, robustness analysis of a given solution is relatively easy

Main takeaways

- While optimization under uncertainty can be difficult, robustness analysis of a given solution is relatively easy
- Our novel method offers many practical advantages over existing methods

Main takeaways

- While optimization under uncertainty can be difficult, robustness analysis of a given solution is relatively easy
- Our novel method offers many practical advantages over existing methods
  - Simple & accessible

Main takeaways

- While optimization under uncertainty can be difficult, robustness analysis of a given solution is relatively easy
- Our novel method offers many practical advantages over existing methods
  - Simple & accessible
  - Applicable to a wide variety of problems

Main takeaways

- While optimization under uncertainty can be difficult, robustness analysis of a given solution is relatively easy
- Our novel method offers many practical advantages over existing methods
  - Simple & accessible
  - Applicable to a wide variety of problems
  - Computationally efficient

Main takeaways

- While optimization under uncertainty can be difficult, robustness analysis of a given solution is relatively easy
- Our novel method offers many practical advantages over existing methods
  - Simple & accessible
  - Applicable to a wide variety of problems
  - Computationally efficient
- Numerical results have been very promising!

Thanks for listening!

If interested in the paper, please contact us at: `j.s.starreveld@uva.nl`

References

Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Math. Opre. Res*, *23*(4), 769–805.

Bertsimas, D., Gupta, V., & Kallus, N. (2018). Data-driven robust optimization. *Mathematical Programming*, *167*(2), 235–292.

Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations research*, *52*(1), 35–53.

Calafiore, G. C. (2016). Repetitive scenario design. *IEEE Transactions on Automatic Control*, *62*(3), 1125–1137.

References (cont.)

Calafiore, G. C., & Campi, M. C. (2005). Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, *102*(1), 25–46.

Campi, M. C., & Garatti, S. (2008). The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, *19*(3), 1211–1230.

Carè, A., Garatti, S., & Campi, M. C. (2014). Fast—fast algorithm for the scenario technique. *Operations Research*, *62*(3), 662–671.

Charnes, A., & Cooper, W. (1959). Chance-constrained programming. *Management Science*, *6*(1), 73–79.

Garatti, S., Campi, M., et al. (2022). Complexity is an effective observable to tune early stopping in scenario optimization. *IEEE Transactions on Automatic Control*.

References (cont.)

Shapiro, A., & Nemirovski, A. (2005). On complexity of stochastic programming problems. *Continuous optimization: Current trends and modern applications*, 111–146.

Vayanos, P., Kuhn, D., & Rustem, B. (2012). A constraint sampling approach for multi-stage robust optimization. *Automatica*, *48*(3), 459–471.

Yanıkoğlu, İ., & den Hertog, D. (2013). Safe approximations of ambiguous chance constraints using historical data. *INFORMS Journal on Computing*, *25*(4), 666–681.