Distributionally robust second-order stochastic dominance constrained optimization with Wasserstein ball

Jia Liu

Xi'an Jiaotong University

TEL:086-82663661, E-mail: jialiu@xjtu.edu.cn

Joint work with Yu Mei and Zhiping Chen (Xi'an Jiaotong University)

Outline

- Introduction
- Model description
- Lower bound approximation
- Upper bound approximation
- Numerical results
- Conclusion

Introduction

Basic definitions of stochastic dominance:

Definition 1 (FSD)

$X \in \mathscr{L}_p$ dominates $Y \in \mathscr{L}_p$ in the first order, denoted $X \succeq_{(1)} Y$, if

$$P\{X \le \eta\} \le P\{Y \le \eta\}, \quad \forall \eta \in \mathbb{R}.$$

We define expected shortfall function
$F_2(X;\eta) = \int_{-\infty}^{\eta} F(X;\alpha)d\alpha = \mathbb{E}[(\eta - X)_+]$.

Definition 2 (SSD)

$X \in \mathscr{L}_p$ dominates $Y \in \mathscr{L}_p$ in the second order, denoted $X \succeq_{(2)} Y$, if

$$F_2(X;\eta) \le F_2(Y;\eta), \ \forall \eta \in \mathbb{R}.$$

Second-order stochastic dominance is particularly popular in industry since it models risk-averse preferences.

Introduction

Proposition 1

- $X \succeq_{(1)} Y$ *iff* $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ *for all* $u \in \mathscr{U}_1$, *here* $\mathscr{U}_1$ *denotes the set of all nondecreasing functions* $u : R \to R$.
- $X \succeq_{(2)} Y$ *iff* $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ *for all* $u \in \mathscr{U}_2$, *here* $\mathscr{U}_2$ *denotes the set of all concave and nondecreasing functions* $u : R \to R$.

- Dentcheva and Ruszczyński (2003) first considered optimization problem with SSD and derived the optimality conditions.
- Dentcheva and Ruszczyński (2006) developed duality relations and solved the dual problem by utilizing the piecewise linear structure of the dual functional
- Luedtke (2008) get new linear formulations for SSD with finite distributed benchmark
- Drapkin, Gollmer, Gotzes, Schultz, et al. (2011a,2011b) study cases where the random variables are induced by mixed-integer linear recourse

Introduction

Solution methods

- Sampling approaches are the most popular solution method (see, Dentcheva and Ruszczyński, 2003, Liu, Sun and Xu, 2016)
- Cut plane methods are the most efficient solution algorithm (see, e.g., Rudolf and Ruszczyński, 2003; Homem-de-Mello and Mehrotra, 2009; Sun, Xu, et al., 2013).

Strong application background in finance

- e.g., portfolio selection, index tracking applications (Dentcheva and Ruszczyński, 2006, Meskarian, Fliege and Xu 2014; Chen, Zhuang, L., 2019)

Definition and model description

Definition 3

X dominates Y robustly in the second order over a set of probability measures $\mathcal{Q} \subset \mathscr{P}$, denoted by $X \succeq_{(2)}^{\mathcal{Q}} Y$, if

$$\mathbb{E}_P[u(X)] \geq \mathbb{E}_P[u(Y)], \ \forall u \in \mathcal{U}, \ \forall P \in \mathcal{Q},$$

where $\mathcal{U}$ is the set of all non-decreasing and concave utility functions.

We investigate the following distributionally robust SSD constrained optimization problem

$$(P_{SSD}) \qquad \begin{aligned} \min_{z \in Z} \quad & f(z) \\ \text{s.t.} \quad & z^T\xi \succeq_{(2)}^{\mathcal{Q}} z_0^T\xi. \end{aligned}$$

Problem (P_{SSD}) can be rewritten as

$$\begin{aligned} \min_{z \in Z} \quad & f(z) \\ \text{s.t.} \quad & \mathbb{E}_P[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+] \leq 0, \ \forall \eta \in \mathbb{R}, \ \forall P \in \mathcal{Q}. \end{aligned} \tag{1}$$

Assumption

Assumption 1

Ξ is polyhedral, i.e., $\Xi = \{\xi \in \mathbb{R}^n \mid C\xi \leq d\}$, where $C \in \mathbb{R}^{l\times n}$, $d \in \mathbb{R}^l$, and $z_0^T\Xi := \{z_0^T\xi \mid \xi \in \Xi\}$ is a compact set.

Problem (1) can be formulated as

$$\begin{aligned} \min_{z\in Z} \quad & f(z) \\ \text{s.t.} \quad & \mathbb{E}_P[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+] \leq 0, \ \forall \eta \in \mathcal{R} := z_0^T\Xi, \ \forall P \in \mathcal{Q}. \end{aligned} \tag{2}$$

Introduction

Ambiguity sets in DRO

- Moment-based ambiguity sets (see, Zymler, Kuhn and Rustem, 2013)
- Distance-based ambiguity sets
  - Kullback-Leibler divergence (see, Hu, Hong, 2014; L., Lisser and Chen, 2019)
  - ϕ-divergence (see, Jiang and Guan, 2016)
  - Wasserstein distance (see, Gao and Kleywegt, 2016, Esfahani and Kuhn,2015,2018, Chen, Kuhn and Wiesemann, 2018, Ji and Lejeune, 2020, Xie, 2021)
- Mixture distribution ambiguity sets... (see Zhu et al. 2014, Chen, Peng, L., 2018)

Related works

Distributional robustness + Stochastic dominance

- Dentcheva and Ruszczyński (2010) introduced distributionally robust SD and establish the optimality conditions.
- Dupačová and Kopa (2014) modeled the ambiguity of the distribution by a linear combination of a nominal distribution and a known contamination distribution with the combination parameter being in a parametric uncertainty set.
- Guo, Xu and Zhang (2017) proposed a discrete approximation scheme for the moment-based ambiguity sets and approximately solved the resulting stochastic optimization problem with distributionally robust SSD constraints.
- Chen and Jiang (2018) and Zhang et al. (2021) studied stability of DRO problems with kth order SD constraints induced by full random recourse.

Data-driven Wasserstein ambiguity set

Our motivation: Distributionally robust SSD + Wasserstein ball

Definition 4

The *Kantorovich metric* $d_K : \mathscr{P}(\Xi) \times \mathscr{P}(\Xi) \to \mathbb{R}_+$ is defined via

$$d_K(P,Q) := \inf_{\pi} \left\{ \int_{\Xi^2} \|\xi_1 - \xi_2\| \pi(d\xi_1, d\xi_2) : \begin{array}{l} \pi \text{ is a joint distribution of } \xi_1 \text{ and } \xi_2 \\ \text{with marginals } P \text{ and } Q, \text{ respectively} \end{array} \right\}.$$

Given N observations $\{\widehat{\xi}_i\}_{i=1}^N$ of ξ, we define the data-driven Wasserstein ambiguity set $\mathcal{Q}$ as a ball centered at the empirical distribution $\widehat{P}_N = \frac{1}{N}\sum_{i=1}^N \delta_{\widehat{\xi}_i}$,

$$\mathcal{Q} := \{P \in \mathscr{P}(\Xi) : d_K(P, \widehat{P}_N) \le \epsilon\}. \tag{3}$$

Main difficulties

- The semi-infiniteness induced from both the SSD and the distributionally robust counterpart are the main challenge.
- Distributionally robust SSD constraints are non-smooth such that gradient based methods fail to work here.
- Compared to moment-based ambiguity sets, the Wasserstein distance contains an extra optimization problem on computing the optimal transportation from the true distribution to the nominal distribution. Such an inner-level optimization problem leads a min-max-min structure and non-convexity of the distributionally robust SSD constraints.

Contributions

- Lower bound approximation:
  - use sample approximation approach
  - establish the quantitative convergency of the approximation
- Upper bound approximation:
  - propose a novel split-and-dual decomposition framework
  - prove convergency of approximation approach and quantitatively estimate the approximation error
- Peng and Delage (2020) find lower bound by sampling approach; upper bound by reformulation as a multistage robust optimization problem.

Strong duality

Gao and Kleywegt (2016)

Lemma 5

If $\Psi(\xi)$ *is proper, continuous, and for some* $\zeta \in \Xi$*, the growth rate* $\kappa := \limsup_{\|\xi-\zeta\|\to\infty} \frac{\Psi(\xi)-\Psi(\zeta)}{\|\xi-\zeta\|} < \infty$*, then the optimal values of*

$$\sup_{P\in\mathscr{P}(\Xi)} \left\{ \int_{\Xi} \Psi(\xi)P(d\xi) : d_K(P, \widehat{P}_N) \le \epsilon \right\}$$

and

$$\min_{\lambda\ge 0} \left\{ \lambda\epsilon + \frac{1}{N}\sum_{i=1}^{N} \sup_{\xi\in\Xi}[\Psi(\xi) - \lambda\|\xi - \widehat{\xi}_i\|] \right\} \qquad (4)$$

are equal. Moreover, the optimal solution set of (4) *is nonempty and compact.*

Lower bound approximation

Let $\Xi_{\mathcal{N}} = \{\bar{\xi}_j\}_{j=1}^{\mathcal{N}}$ be a set of finite samples in Ξ and $\varGamma_{\mathcal{M}} = \{\eta_k\}_{k=1}^{\mathcal{M}}$ be a set of finite samples in $\mathcal{R} = [\mathcal{R}_{\min}, \mathcal{R}_{\max}]$.

We then approximate the ambiguity set $\mathcal{Q}$ by the following Wasserstein ball:

$$\mathcal{Q}_{\mathcal{N}} := \{P \in \mathscr{P}(\Xi_{\mathcal{N}}) : d_K(P, \widehat{P}_{\mathcal{N}}) \leq \epsilon\}.$$

We have a lower bound approximation of problem (2):

$$\begin{aligned} \min_{z \in Z} \quad & f(z) \\ \text{s.t.} \quad & \mathbb{E}_P[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+] \leq 0, \ \forall \eta \in \varGamma_{\mathcal{M}}, \ \forall P \in \mathcal{Q}_{\mathcal{N}}. \end{aligned} \tag{5}$$

Tractability of the lower bound approximation

By Lemma 5 and introducing auxiliary variables, we have a linear programming reformulation of problem (5)

$$
(P_{SSD-L}) \quad
\begin{aligned}
\min_{z,\lambda,\beta,s} \quad & f(z) \\
\text{s.t.} \quad & \lambda_k \epsilon - \sum_{i=1}^{N} \frac{1}{N} \beta_{ik} \le 0, \ k = 1, \cdots, \mathcal{M}, \\
& \beta_{ik} + s_{jk} \le \lambda_k \|\bar{\xi}_j - \widehat{\xi}_i\| + (\eta_k - z_0^T \bar{\xi}_j)_+, \\
& \qquad i = 1, \cdots, N, \ j = 1, \cdots, \mathcal{N}, \ k = 1, \cdots, \mathcal{M}, \\
& s_{jk} \ge \eta_k - z^T \bar{\xi}_j, \ j = 1, \cdots, \mathcal{N}, \ k = 1, \cdots, \mathcal{M}, \\
& z \in Z, \lambda \in \mathbb{R}_+^{\mathcal{M}}, \beta \in \mathbb{R}^{N \times \mathcal{M}}, s \in \mathbb{R}_+^{\mathcal{N} \times \mathcal{M}}.
\end{aligned}
$$

Quantitative analysis of the lower approximation

Assumption 2

There exist a point $\bar{z} \in Z$ *and a constant* $\theta > 0$ *such that*

$$\sup_{\eta \in \mathcal{R}} \sup_{P \in \mathcal{Q}} \mathbb{E}_P[(\eta - \bar{z}^T \xi)_+ - (\eta - z_0^T \xi)_+] < -\theta;$$

$\lim_{\mathcal{N} \to \infty} \alpha_{\mathcal{N}} = 0$ *and* $\lim_{\mathcal{M} \to \infty} \gamma_{\mathcal{M}} = 0$, *where* $\alpha_{\mathcal{N}} := \sup_{\xi \in \Xi} \inf_{\xi' \in \Xi_{\mathcal{N}}} \|\xi - \xi'\|$ *and* $\Gamma_{\mathcal{M}}$ *by* $\gamma_{\mathcal{M}} := \sup_{\eta \in \mathcal{R}} \inf_{\eta' \in \Gamma_{\mathcal{M}}} |\eta - \eta'|$.

Quantitative analysis of the lower approximation

Theorem 6

Given Assumption 2, the following assertions hold.
(i) For any $\mathcal{N}$ and $\mathcal{M}$,

$$\mathbb{H}(\mathcal{F}_{\mathcal{N},\mathcal{M}},\mathcal{F}) \le \frac{2D_Z}{\theta}(L\alpha_{\mathcal{N}}+\gamma_{\mathcal{M}}).$$

(ii) $\lim_{\substack{\mathcal{N}\to\infty,\\ \mathcal{M}\to\infty}} v_{\mathcal{N},\mathcal{M}} = v$ *and* $\limsup_{\substack{\mathcal{N}\to\infty,\\ \mathcal{M}\to\infty}} \mathcal{S}_{\mathcal{N},\mathcal{M}} \subset \mathcal{S}$.
(iii) If, in addition, the objective function f is Lipschitz continuous with modulus L_f, then for any $\mathcal{N}$ and $\mathcal{M}$,

$$|v_{\mathcal{N},\mathcal{M}} - v| \le \frac{2D_Z L_f}{\theta}(L\alpha_{\mathcal{N}}+\gamma_{\mathcal{M}}).$$

Moreover, if there exists a positive constant ρ such that $f(z)-v \ge \rho d(z,\mathcal{S})^2, \forall z\in\mathcal{F}$, *then for sufficiently large $\mathcal{N}$ and $\mathcal{M}$,*

$$\mathbb{D}(\mathcal{S}_{\mathcal{N},\mathcal{M}},\mathcal{S}) \le \Big(\sqrt{\frac{2D_Z}{\theta}} + \sqrt{\frac{4L_f D_Z}{\rho\theta}}\Big)\sqrt{L\alpha_{\mathcal{N}}+\gamma_{\mathcal{M}}}. \tag{6}$$

Cutting-plane method

Algorithm 3.1 Cutting-plane Method

Start from $\iota = 1$ and $\mathcal{J}_1^\iota = \mathcal{J}_2^\iota = \emptyset$.
while $\iota \geq 1$ **do**

Solve the approximate problem:

$$
\begin{aligned}
&\min_{z,\lambda,\beta,s} \; f(z)\\
&\text{s.t. } \; \beta_{ik} + s_{jk} \leq \lambda_k \|\bar{\xi}_j - \widehat{\xi}_i\| + (\eta_k - z_0^T \bar{\xi}_j)_+, i = 1, \cdots, N, j \in \mathcal{J}_1^\iota, k \in \mathcal{J}_2^\iota,\\
&\qquad s_{jk} \geq \eta_k - z^T \bar{\xi}_j, j \in \mathcal{J}_1^\iota, k \in \mathcal{J}_2^{l\iota},\\
&\qquad (3.8b), \; z \in Z, \lambda \in \mathbb{R}_+^M, \beta \in \mathbb{R}^{N\times M}, s \in \mathbb{R}_+^{N\times M}.
\end{aligned}
\tag{3.9}
$$

Let $(z^\iota, \lambda^\iota, \beta^\iota, s^\iota)$ denote the optimal solution of problem (3.9).
Calculate

$$
\delta^\iota := \max_{\substack{i\in\{1,\cdots,N\}, j\in\{1,\cdots,N\},\\ k\in\{1,\cdots,M\}}} \left\{ \beta_{ik}^\iota - \lambda_k^\iota \|\bar{\xi}_j - \widehat{\xi}_i\| + (\eta_k - (z^\iota)^T \bar{\xi}_j)_+ - (\eta_k - z_0^T \bar{\xi}_j)_+ \right\}.
$$

if $\delta^\iota \leq 0$ **then**
Stop.
else
Determine

$$
(i^\iota, j^\iota, k^\iota) \in \operatorname*{argmax}_{\substack{i\in\{1,\cdots,N\}, j\in\{1,\cdots,N\},\\ k\in\{1,\cdots,M\}}} \left\{ \beta_{ik}^\iota - \lambda_k^\iota \|\bar{\xi}_j - \widehat{\xi}_i\| + (\eta_k - (z^\iota)^T \bar{\xi}_j)_+ - (\eta_k - z_0^T \bar{\xi}_j)_+ \right\}.
$$

Let $\mathcal{J}_1^{\iota+1} = \mathcal{J}_1^\iota \cup j^\iota$, $\mathcal{J}_2^{\iota+1} = \mathcal{J}_2^\iota \cup k^\iota$ and $\iota \leftarrow \iota + 1$.
end if
end while

Upper bound approximation

Notice that problem (2) can be rewritten as

$$\begin{array}{ll} \min_{z\in Z} & f(z) \\ \text{s.t.} & \sup_{P\in\mathcal{Q}} \sup_{\eta\in\mathcal{R}} \mathbb{E}_P[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+] \le 0. \end{array} \tag{7}$$

If we exchange the order of operators $\sup_{\eta\in\mathcal{R}}$ and $\mathbb{E}_P$ in problem (7), we obtain an upper bound approximation for problem (7). However, such an upper bound approximation might be loose or even infeasible since the gap

$$\mathbb{E}_P\left[\sup_{\eta\in\mathcal{R}}\{(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+\}\right] - \sup_{\eta\in\mathcal{R}} \mathbb{E}_P[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+] \tag{8}$$

might be large. The larger the range $\mathcal{R}$ of η, the larger the gap in (8). As an extreme case, when $\mathcal{R}$ reduces to a singleton, the gap (8) becomes 0.

Split-and-dual framework

In detail, we divide $\mathcal{R} = [\mathcal{R}_{\min}, \mathcal{R}_{\max}]$ into $\mathcal{K}$ intervals with disjoint interiors, $[\underline{\eta}_k, \bar{\eta}_k]$, $k = 1, \cdots, \mathcal{K}$, where the boundary points of the intervals are specified by $\underline{\eta}_k = \mathcal{R}_{\min} + (k-1)\frac{\mathcal{R}_{\max} - \mathcal{R}_{\min}}{\mathcal{K}}$, $\bar{\eta}_k = \mathcal{R}_{\min} + k\frac{\mathcal{R}_{\max} - \mathcal{R}_{\min}}{\mathcal{K}}$, $k = 1, \cdots, \mathcal{K}$. Problem (7) can be reformulated as

$$\begin{aligned} \min_{z \in Z} \quad & f(z) \\ \text{s.t.} \quad & \sup_{P \in \mathcal{Q}} \max_{1 \le k \le \mathcal{K}} \sup_{\eta \in [\underline{\eta}_k, \bar{\eta}_k]} \mathbb{E}_P[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+] \le 0, \end{aligned}$$

Exchanging the order of operators $\sup_{\eta \in [\underline{\eta}_k, \bar{\eta}_k]}$ and $\mathbb{E}_P$, we have the following approximation problem

$$\begin{aligned} \min_{z \in Z} \quad & f(z) \\ \text{s.t.} \quad & \sup_{P \in \mathcal{Q}} \mathbb{E}_P\Big[\sup_{\eta \in [\underline{\eta}_k, \bar{\eta}_k]} \{(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+\}\Big] \le 0, \ k = 1, \cdots, \mathcal{K}. \end{aligned} \tag{9}$$

Reformulation of the upper approximation

By applying Lemma 5, we have a reformulation of problem (9)

$$\min_{z \in Z, \lambda \in \mathbb{R}_+^{\mathcal{K}}} \quad f(z) \tag{10a}$$

$$\text{s.t.} \quad \lambda_k \epsilon + \frac{1}{N} \sum_{i=1}^{N} \sup_{\xi \in \Xi} \Bigg\{ \tag{10b}$$

$$\sup_{\eta \in [\underline{\eta}_k, \bar{\eta}_k]} \{(\eta - z^T \xi)_+ - (\eta - z_0^T \xi)_+\} - \lambda_k \|\xi - \widehat{\xi}_i\| \Bigg\} \le 0, \ k = 1, \cdots, \mathcal{K}.$$

We write (10b) as

$$\lambda_k \epsilon + \frac{1}{N} \sum_{i=1}^{N} V_S^{ik} \le 0, \ k = 1, \cdots, \mathcal{K}, \tag{11}$$

where

$$V_S^{ik} := \sup_{(\xi, \eta) \in \Xi \times [\underline{\eta}_k, \bar{\eta}_k]} (\eta - z^T \xi)_+ - (\eta - z_0^T \xi)_+ - \lambda_k \|\xi - \widehat{\xi}_i\|, \ i = 1, \cdots, N, \ k = 1, \cdots, \mathcal{K}.$$

We can split the problem into two convex sub-problems:

$$(P^{ik}_{SSD-1}) \quad \begin{aligned} V^{ik}_{S1} = \sup_{\xi,\eta,s,m} \quad & \eta - z^T\xi - s - \lambda_k m \\ \text{s.t.} \quad & s \ge \eta - z_0^T \xi, \\ & \eta - z^T\xi \ge 0, \\ & C\xi \le d, \\ & s \ge 0, \\ & \eta \ge \underline{\eta}_k, \\ & \eta \le \bar{\eta}_k, \\ & \|\xi - \widehat{\xi_i}\| \le m. \end{aligned}$$

$$(P^{ik}_{SSD-2}) \quad \begin{aligned} V^{ik}_{S2} = \sup_{\xi,\eta,s,m} \quad & -s - \lambda_k m \\ \text{s.t.} \quad & s \ge \eta - z_0^T \xi, \\ & \eta - z^T\xi \le 0, \\ & C\xi \le d, \\ & s \ge 0, \\ & \eta \ge \underline{\eta}_k, \\ & \eta \le \bar{\eta}_k, \\ & \|\xi - \widehat{\xi_i}\| \le m. \end{aligned}$$

And we have

$$V^{ik}_S = \max\{V^{ik}_{S1}, V^{ik}_{S2}\}. \tag{12}$$

Conic duality

$$
\begin{aligned}
\tilde{V}_{S1}^{ik} &= \inf_{\mu,\nu} d^T\nu - \widehat{\xi}_i^T(z - \mu_1 z_0 + \mu_2 z + C^T\nu) - \mu_3 \underline{\eta}_k + (1-\mu_1 + \mu_2 + \mu_3)\bar{\eta}_k \\
(D_{SSD-1}^{ik}) \quad & \text{s.t. } \mu_1 \le 1,\ 1 - \mu_1 + \mu_2 + \mu_3 \ge 0, \\
& \|z - \mu_1 z_0 + \mu_2 z + C^T\nu\| \le \lambda_k, \\
& \mu \in \mathbb{R}_+^3,\ \nu \in \mathbb{R}_+^l.
\end{aligned}
$$

$$
\begin{aligned}
\tilde{V}_{S2}^{ik} &= \inf_{\mu,\nu} d^T\nu - \widehat{\xi}_i^T(-\mu_1 z_0 - \mu_2 z + C^T\nu) - \mu_3 \underline{\eta}_k + (-\mu_1 - \mu_2 + \mu_3)\bar{\eta}_k \\
(D_{SSD-2}^{ik}) \quad & \text{s.t. } \mu_1 \le 1,\ -\mu_1 - \mu_2 + \mu_3 \ge 0, \\
& \|-\mu_1 z_0 - \mu_2 z + C^T\nu\| \le \lambda_k, \\
& \mu \in \mathbb{R}_+^3,\ \nu \in \mathbb{R}_+^l.
\end{aligned}
$$

Upper bound

Theorem 7

Given Assumption 1, the optimal value of the following optimization problem

$$
\begin{aligned}
&\min \quad f(z)\\
&\text{s.t.} \quad \lambda_k \epsilon + \frac{1}{N}\sum_{i=1}^{N} V^{ik} \le 0,\ k = 1,\cdots,\mathcal{K},\\
&(P_{SSD-U}) \quad \left.\begin{array}{l}
\mu_1^{ik} \le 1,\ \tilde{\mu}_1^{ik} \le 1,\ 1-\mu_1^{ik}+\mu_2^{ik}+\mu_3^{ik} \ge 0,\ -\tilde{\mu}_1^{ik}-\tilde{\mu}_2^{ik}+\tilde{\mu}_3^{ik} \ge 0,\\
V^{ik} \ge d^T\nu^{ik} - \widehat{\xi}_i^T(z-\mu_1^{ik}z_0+\mu_2^{ik}z+C^T\nu^{ik}) - \mu_3^{ik}\underline{\eta}_k + (1-\mu_1^{ik}+\mu_2^{ik}+\mu_3^{ik})\bar{\eta}_k,\\
V^{ik} \ge d^T\tilde{\nu}^{ik} - \widehat{\xi}_i^T(-\tilde{\mu}_1^{ik}z_0-\tilde{\mu}_2^{ik}z+C^T\tilde{\nu}^{ik}) - \tilde{\mu}_3^{ik}\underline{\eta}_k + (-\tilde{\mu}_1^{ik}-\tilde{\mu}_2^{ik}+\tilde{\mu}_3^{ik})\bar{\eta}_k,\\
\|z-\mu_1^{ik}z_0+\mu_2^{ik}z+C^T\nu^{ik}\| \le \lambda_k,\ \|-\tilde{\mu}_1^{ik}z_0-\tilde{\mu}_2^{ik}z+C^T\tilde{\nu}^{ik}\| \le \lambda_k,\\
\mu^{ik}\in\mathbb{R}_+^3, \nu^{ik}\in\mathbb{R}_+^l, \tilde{\mu}^{ik}\in\mathbb{R}_+^3, \tilde{\nu}^{ik}\in\mathbb{R}_+^l, V^{ik}\in\mathbb{R},
\end{array}\right\}\\
&\qquad\qquad\qquad\qquad i = 1,\cdots,N,\ k = 1,\cdots,\mathcal{K},\\
&\quad z\in Z, \lambda\in\mathbb{R}_+^{\mathcal{K}}.
\end{aligned}
$$

is an upper bound to that of problem (P_{SSD}).

Quantitative analysis of the upper approximation

Let

$$g(z,\mathcal{K}) := \max_{1\le k\le \mathcal{K}} \sup_{P\in\mathcal{Q}} \mathbb{E}_P\Big[\sup_{\eta\in[\underline{\eta}_k,\bar{\eta}_k]} \Big\{(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+\Big\}\Big],$$

$$\begin{aligned} g(z) :&= \sup_{P\in\mathcal{Q}} \sup_{\eta\in\mathcal{R}} \mathbb{E}_P\left[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+\right] \\ &= \max_{1\le k\le \mathcal{K}} \sup_{P\in\mathcal{Q}} \sup_{\eta\in[\underline{\eta}_k,\bar{\eta}_k]} \mathbb{E}_P\left[(\eta - z^T\xi)_+ - (\eta - z_0^T\xi)_+\right]. \end{aligned}$$

Proposition 2

Given Assumption 1, for any positive integer $\mathcal{K}$, $g(\cdot,\mathcal{K})$ and $g(\cdot)$ are Lipschitz continuous with modulus $C = \sup_{P\in\mathcal{Q}} \mathbb{E}_P[\|\xi\|] < \infty$.

Proposition 3

Given Assumption 1, we have that $g(z,\mathcal{K}) - g(z) \le 2\frac{\mathcal{R}_{\max}-\mathcal{R}_{\min}}{\mathcal{K}}$, and $\lim_{\mathcal{K}\to\infty} g(z,\mathcal{K}) = g(z)$, uniformly with respect to $z\in Z$.

Non-differentiable MFCQ

Definition 8

(ND-MFCQ) Let $F(t) := \{x \in \mathbb{R}^n \mid g_j(x,t) \leq 0, j \in J\}$ with subdifferentiable g_j, here t is the parameter in the constraints. Given $\bar{t}$ and $\bar{x} \in F(\bar{t})$, if there exist some vector θ and real constants $\sigma < 0$, $\alpha_1 > 0$, $\alpha_2 > 0$ such that

$$\langle \varsigma, \theta \rangle \leq \sigma < 0, \ \forall \varsigma \in \partial g_j(x,t), \forall x : \|x - \bar{x}\| \leq \alpha_1, \forall t : \|t - \bar{t}\| \leq \alpha_2, \forall j \in J_0(\bar{x}, \bar{t}),$$

where $J_0(\bar{x}, \bar{t}) := \{j \in J \mid g_j(\bar{x}, \bar{t}) = 0\}$, then we say that non-differentiable MFCQ (ND-MFCQ) holds at $(\bar{x}, \bar{t})$ with θ, σ, α_1 and α_2,

ND-MFCQ is equivalent to MFCQ if differentiable

Quantitative analysis of the upper approximation

Assumption 3

The optimal solution set of problem (9) *with* $\mathcal{K} = 1$, *denoted by* $\mathcal{S}_1$, *is nonempty.*

Theorem 9

Given Assumptions 1 and 3. For some $z^* \in \mathcal{S}$, *assume that ND-MFCQ holds at* $(z^*, 0)$ *with* θ, σ, α_1, *and* α_2 *as is defined in Definition 8. If the objective function* f *is Lipschitz continuous with modulus* L_f, *then for* $\mathcal{K} \geq \max\left\{\frac{1}{\alpha_2}, \frac{2}{|\sigma|}\frac{\mathcal{R}_{\max}-\mathcal{R}_{\min}}{\alpha_1}||\theta||, -2\frac{\mathcal{R}_{\max}-\mathcal{R}_{\min}}{g(z^*)}\left(C\frac{||\theta||}{|\sigma|}+1\right)\right\}$, *we have that*

$$|v_{\mathcal{K}} - v| \leq L_f \frac{2||\theta||}{|\sigma|}\frac{\mathcal{R}_{\max} - \mathcal{R}_{\min}}{\mathcal{K}},$$

and $\lim_{\mathcal{K}\to\infty} v_{\mathcal{K}} = v$.

Upper bound

Theorem 10

Given Assumption 1, the optimal value of the following optimization problem

$$
\begin{aligned}
&\min \quad f(z)\\
&\text{s.t.} \quad \lambda_k \epsilon + \frac{1}{N}\sum_{i=1}^{N} V^{ik} \le 0,\ k = 1,\cdots,\mathcal{K},\\
&(P_{SSD-U}) \quad \left.\begin{array}{l}
\mu_1^{ik} \le 1,\ \tilde{\mu}_1^{ik} \le 1,\ 1-\mu_1^{ik}+\mu_2^{ik}+\mu_3^{ik} \ge 0,\ -\tilde{\mu}_1^{ik}-\tilde{\mu}_2^{ik}+\tilde{\mu}_3^{ik} \ge 0,\\
V^{ik} \ge d^T\nu^{ik} - \widehat{\xi}_i^T(z-\mu_1^{ik}z_0+\mu_2^{ik}z+C^T\nu^{ik}) - \mu_3^{ik}\underline{\eta}_k + (1-\mu_1^{ik}+\mu_2^{ik}+\mu_3^{ik})\bar{\eta}_k,\\
V^{ik} \ge d^T\tilde{\nu}^{ik} - \widehat{\xi}_i^T(-\tilde{\mu}_1^{ik}z_0-\tilde{\mu}_2^{ik}z+C^T\tilde{\nu}^{ik}) - \tilde{\mu}_3^{ik}\underline{\eta}_k + (-\tilde{\mu}_1^{ik}-\tilde{\mu}_2^{ik}+\tilde{\mu}_3^{ik})\bar{\eta}_k,\\
\|z-\mu_1^{ik}z_0+\mu_2^{ik}z+C^T\nu^{ik}\| \le \lambda_k,\ \|-\tilde{\mu}_1^{ik}z_0-\tilde{\mu}_2^{ik}z+C^T\tilde{\nu}^{ik}\| \le \lambda_k,\\
\mu^{ik}\in\mathbb{R}_+^3, \nu^{ik}\in\mathbb{R}_+^l, \tilde{\mu}^{ik}\in\mathbb{R}_+^3, \tilde{\nu}^{ik}\in\mathbb{R}_+^l, V^{ik}\in\mathbb{R},
\end{array}\right\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1,\cdots,N,\ k = 1,\cdots,\mathcal{K},\\
&\quad z\in Z, \lambda\in\mathbb{R}_+^{\mathcal{K}}.
\end{aligned}
$$

is an upper bound to that of problem (P_{SSD}).

Sequential convex approximation

Algorithm 4.1 Sequential convex approximation

Start from $z^\iota \in Z$, $\iota = 1$.
while $\iota \geq 1$ **do**
  Solve problem (P_{SSD-U}) with an additional constraint $z = z^\iota$. Denote the optimal $\mu, \tilde{\mu}$ by $\mu^\iota, \tilde{\mu}^\iota$, respectively.
  Solve problem (P_{SSD-U}) with additional constraints $\mu = \mu^\iota, \tilde{\mu} = \tilde{\mu}^\iota$. Denote the optimal z by $z^{\iota+1}$.
  if $z^{\iota+1} = z^\iota$ **then**
    Break.
  else
    $\iota \leftarrow \iota + 1$.
  end if
end while

Numerical results

Table: The optimal values and the optimal solutions of the lower and upper bound approximations

lower bound approximation ((P_{SSD-L}) or Algorithm 3.1)				upper bound approximation (Algorithm 4.1)			Gap
$\mathcal{N}$	$\mathcal{M}$	Optimal value	Optimal solution	$\mathcal{K}$	Optimal value	Optimal solution	
100	100	0.2922	$(0.4229, 0.4027)^T$	10	0.4097	$(0.8010, 0.1564)^T$	40.2122%
200	200	0.2964	$(0.4266, 0.4077)^T$	11	0.3044	$(0.4590, 0.4002)^T$	2.6991%
300	300	0.3014	$(0.4423, 0.4014)^T$	12	0.3025	$(0.4653, 0.3868)^T$	0.3650%

Numerical results

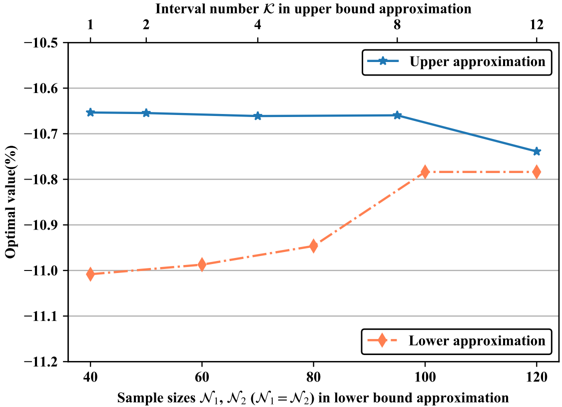

Figure: Optimal values of the lower bound approximation with respect to $\mathcal{N}, \mathcal{M}$ and that of the upper bound approximation with respect to $\mathcal{K}$.

Numerical results

Table: Optimal values of the lower and upper bound approximations, and their relative gaps with respect to different robust radii.

Robust radius ϵ	Optimal values (%) lower bound approximation	Optimal values (%) upper bound approximation	Gap
10^{-5}	-10.8775	-10.8268	0.4661%
10^{-4}	-10.7838	-10.7389	0.4164%
10^{-3}	-10.7836	-10.6536	1.2055%
10^{-2}	-10.7823	-10.6535	1.1946%
0.1	-10.7689	-10.6534	1.0725%
0.5	-10.6885	-10.6534	0.3284%
1	-10.6534	-10.6534	0%

Conclusions

Summary:

- Study a distributionally robust SSD constrained optimization problem
- Adopt the sample approximation approach to develop a linear programming formulation to obtain a lower bound approximation
- Establish the quantitative convergency for the lower bound approximation problem
- Propose a novel split-and-dual decomposition framework to derive an upper bound approximation
- Quantitatively estimate the approximation error between the optimal value of the upper bound approximation and that of the original problem

Further works:

- Modifying the design of cutting-planes
- Investigate the critical number of intervals for enhancing the practicality of the upper bound approximation
- Consider distributionally robust multivariate robust SSD constrained optimization problem

Thank you!