On Distributionally Robust Chance Constrained Program with Wasserstein Distance

Weijun Xie

ISE, Virginia Tech

Mathematical Optimization of Systems Impacted by Rare, High-Impact Random Events, Jun 24 - 28, 2019

Distributionally Robust Chance Constrained Program (DRCCP)

Consider DRCCP as

$$
\begin{aligned}
v^* = \quad & \min_{x} \quad c^\top x && \text{(objective function)} \\
& \text{s.t.} \quad x \in S && \text{(deterministic constraints) e.g., nonnegativity} \\
& \qquad \tilde{A}x \geq \tilde{b} && \text{(uncertain inequalities)}
\end{aligned}
$$

Distributionally Robust Chance Constrained Program (DRCCP)

Consider DRCCP as

$$v^* = \min_{x} \quad c^\top x \qquad \text{(objective function)}$$

$$\text{s.t.} \quad x \in S \qquad \text{(deterministic constraints) e.g., nonnegativity}$$

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\{\tilde{A}x \geq \tilde{b}\} \geq 1-\epsilon \qquad \text{(chance constraint)}$$

where

- $\epsilon \in (0,1)$ is risk parameter
- "Ambiguity Set" $\mathcal{P}$ = a family of probability distributions

Distributionally Robust Chance Constrained Program (DRCCP)

Consider DRCCP as

$$v^* = \min_x \quad c^\top x \qquad \text{(objective function)}$$

$$\text{s.t.} \quad x \in S \qquad \text{(deterministic constraints) e.g., nonnegativity}$$

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\begin{Bmatrix} \tilde{a}_1^\top x \geq \tilde{b}_1 \\ \vdots \\ \tilde{a}_m^\top x \geq \tilde{b}_m \end{Bmatrix} \geq 1-\epsilon \qquad \text{(chance constraint)}$$

where

- $\epsilon \in (0, 1)$ is risk parameter
- "Ambiguity Set" $\mathcal{P}$ = a family of probability distributions
- $m = 1$: single DRCCP; $m > 1$: joint DRCCP

Wasserstein Ambiguity Set

Wasserstein ambiguity set (Esfahani and Kuhn 2015; Zhao and Guan, 2015; Gao and Kleywegt, 2016; Blanchet and Murthy, 2016)

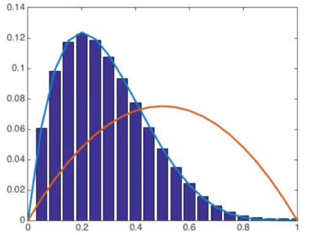
$$\mathcal{P}^W = \left\{ \mathbb{P} : W_q\left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}}\right) \leq \delta \right\},$$

where $W_q\left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}}\right)$ = Wasserstein distance between probability distribution $\mathbb{P}$ and empirical distribution $\mathbb{P}_{\tilde{\zeta}}$.

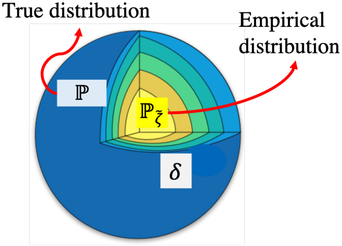
Wasserstein Distance

$$d(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\mathbb{P}} \mathbb{E}_{\mathbb{P}}[\|\tilde{\boldsymbol{x}}_1 - \tilde{\boldsymbol{x}}_2\|]$$

$$(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2) \sim \mathbb{P}, \Pi_{\tilde{\boldsymbol{x}}_1}\mathbb{P} = \mathbb{P}_1, \Pi_{\tilde{\boldsymbol{x}}_2}\mathbb{P} = \mathbb{P}_2$$



Wasserstein Ambiguity



Wasserstein ambiguity set

$$\mathcal{P}_w = \{\mathbb{P} : \mathbb{P}\{\tilde{\xi} \in \Xi\} = 1, W_q(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}}) \leq \delta\}$$

Wasserstein Ambiguity Set

Wasserstein ambiguity set (Esfahani and Kuhn 2015; Zhao and Guan, 2015; Gao and Kleywegt, 2016; Blanchet and Murthy, 2016)

$$\mathcal{P}^W = \left\{ \mathbb{P} : W_q \left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}} \right) \leq \delta \right\},$$

where $W_q \left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}} \right)$ = Wasserstein distance between probability distribution $\mathbb{P}$ and empirical distribution $\mathbb{P}_{\tilde{\zeta}}$.

- Convergence in probability to regular chance constrained program (CCP)
-

DRCCP with Wasserstein Ambiguity Set (DRCCP-W): Existing Works

DRCCP-W set

$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P}\left\{ \tilde{A}x \geq \tilde{b} \right\} \geq 1 - \epsilon \right\},$$

with $\mathcal{P}^W = \left\{ \mathbb{P} : W_q \left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}} \right) \leq \delta \right\}.$

- Hanasusanto et al. (2015) and **X.** and Ahmed (2017) showed that DRCCP-W is a biconvex program.

DRCCP with Wasserstein Ambiguity Set (DRCCP-W): Existing Works

DRCCP-W set

$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P}\left\{ \tilde{A}x \geq \tilde{b} \right\} \geq 1 - \epsilon \right\},$$

with $\mathcal{P}^W = \left\{ \mathbb{P} : W_q \left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}} \right) \leq \delta \right\}$.

- Hanasusanto et al. (2015) and **X.** and Ahmed (2017) showed that DRCCP-W is a biconvex program.
- **X.** and Ahmed (2017) proposed a bicriteria approximation algorithm for a special family of DRCCP-W

DRCCP-W: Summary of Contributions

DRCCP-W set

$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P}\left\{ \tilde{A}x \geq \tilde{b} \right\} \geq 1 - \epsilon \right\}.$$

- DRCCP-W $\equiv$ conditional-value-at-risk (CVaR) constrained optimization
  - Develop inner and outer approximations
- DRCCP-W set Z is mixed integer program representable
  - With big-M coefficients and additional binary variables
- Binary DRCCP-W set (i.e., $S \subseteq \{0,1\}^n$) is submodular constrained
  - Without big-M coefficients and additional binary variables
  - Solvable by Branch and Cut

Outline

- CVaR Reformulation and Related Approximations
- Mixed Integer Program Reformulation
- Binary DRCCP-W and Submodularity
- Concluding Remarks

CVaR Reformulation and Related Approximations

CVaR Reformulation

DRCCP-W set

$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P}\left\{ \tilde{A}x \geq \tilde{b} \right\} \geq 1 - \epsilon \right\},$$

with $\mathcal{P}^W = \left\{ \mathbb{P} : W_q\left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}}\right) \leq \delta \right\}$.

Theorem (Exact Formulation)

$$Z = \left\{ x : \frac{\delta}{\epsilon} + \mathbf{CVaR}_{1-\epsilon}\left[-f(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

where $f(x, \zeta) := \min_{i \in [m]} \inf_{a_i^\top x < b_i} \|(a_i, b_i) - (a_i^\zeta, b_i^\zeta)\|$ and

$\mathbf{CVaR}_{1-\epsilon}\left[\tilde{X}\right] = \min_{\gamma} \left\{ \gamma + \frac{1}{\epsilon} \mathbb{E}_{\mathbb{P}}\left[\tilde{X} - \gamma \right]_+ \right\}.$

CVaR Reformulation

DRCCP-W set

$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P}\left\{ \tilde{A}x \geq \tilde{b} \right\} \geq 1 - \epsilon \right\},$$

with $\mathcal{P}^W = \left\{ \mathbb{P} : W_q\left(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}}\right) \leq \delta \right\}$.

Theorem (Exact Formulation)

$$Z = \left\{ x : \frac{\delta}{\epsilon} + \mathbf{CVaR}_{1-\epsilon}\left[-f(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

where $f(x, \zeta) := \min_{i \in [m]} \inf_{a_i^\top x < b_i} \|(a_i, b_i) - (a_i^\zeta, b_i^\zeta)\|$ and

$\mathbf{CVaR}_{1-\epsilon}\left[\tilde{X}\right] = \min_{\gamma} \left\{ \gamma + \frac{1}{\epsilon} \mathbb{E}_{\mathbb{P}}\left[\tilde{X} - \gamma \right]_+ \right\}.$

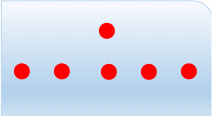
Proof Idea: (1) strong duality of distributionally robust optimization, and (2) break down the indicator function.

CVaR Reformulation: Worst-case Interpretation

Theorem (Exact Formulation)

$$Z = \left\{ x : \frac{\delta}{\epsilon} + \mathbf{CVaR}_{1-\epsilon}\left[-f(x, \tilde{\zeta}) \right] \le 0 \right\},$$

where $f(x, \zeta) := \min_{i \in [m]} \inf_{a_i^\top x < b_i} \|(a_i, b_i) - (a_i^\zeta, b_i^\zeta)\|$



Original empirical samples

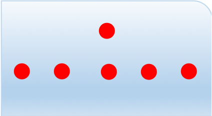
- $N = 6, \epsilon = 1/3$

CVaR Reformulation: Worst-case Interpretation

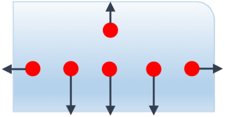
Theorem (Exact Formulation)

$$Z = \left\{ x : \frac{\delta}{\epsilon} + \mathbf{CVaR}_{1-\epsilon}\left[-f(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

where $f(x, \zeta) := \min_{i \in [m]} \inf_{a_i^\top x < b_i} \|(a_i, b_i) - (a_i^\zeta, b_i^\zeta)\|$



Original empirical samples



Moving these samples to boundary of violating constraints

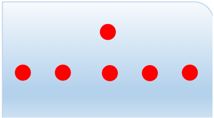
- $N = 6, \epsilon = 1/3$

CVaR Reformulation: Worst-case Interpretation

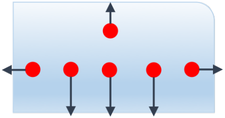
Theorem (Exact Formulation)

$$Z = \left\{ x : \frac{\delta}{\epsilon} + \mathbf{CVaR}_{1-\epsilon}\left[-f(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

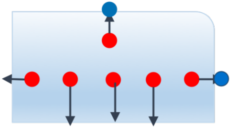
where $f(x,\zeta) := \min_{i\in[m]} \inf_{a_i^\top x < b_i} \|(a_i, b_i) - (a_i^\zeta, b_i^\zeta)\|$



Original empirical samples



Moving these samples to boundary of violating constraints



Due to chance constraint, only limited scenarios can be moved

- $N = 6, \epsilon = 1/3$

CVaR Reformulation: Simplification

DRCCP-W set

$$Z = \left\{ x : \frac{\delta}{\epsilon} + \mathbf{CVaR}_{1-\epsilon}\left[-f(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

where $f(x, \zeta) := \min_{i \in [m]} \inf_{a_i^\top x < b_i} \|(a_i, b_i) - (a_i^\zeta, b_i^\zeta)\|$.

CVaR Reformulation: Simplification

DRCCP-W set

$$Z = \left\{ x : \frac{\delta}{\epsilon} + \mathbf{CVaR}_{1-\epsilon}\left[-f(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

where $f(x, \zeta) := \min_{i \in [m]} \frac{\max\left\{ (a_i^{\zeta})^\top x - b_i^{\zeta}, 0 \right\}}{\|(x, 1)\|_*}$.

CVaR Reformulation: Simplification

DRCCP-W set

$$Z = \left\{ x : \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta}) \right] \le 0 \right\},$$

where $\widehat{f}(x,\zeta) := \min_{i\in[m]} \max\left\{ (a_i^\zeta)^\top x - b_i^\zeta, 0 \right\}.$

- By positive homogeneity of coherent risk measures

CVaR Reformulation: Simplification

DRCCP-W set

$$Z = \left\{ x : \frac{\delta}{\epsilon} \|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta}) \right] \leq 0 \right\},$$

where $\widehat{f}(x,\zeta) := \max\left\{ \min_{i\in[m]} \left[(a_i^\zeta)^\top x - b_i^\zeta \right], 0 \right\}.$

- By positive homogeneity of coherent risk measures
- Switch minimax to maximin

Outer Approximation

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \geq \mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right) := \min\left\{s : F_{\tilde{X}}(s) \geq 1-\epsilon\right\}.$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right)$.

Theorem (Outer Approximation)

$$Z = \left\{x : \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta})\right] \leq 0\right\}$$

where $\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right], 0\right\}$.

Outer Approximation

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \geq \mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right) := \min\left\{s : F_{\tilde{X}}(s) \geq 1-\epsilon\right\}.$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right)$.

Theorem (Outer Approximation)

$$Z \subseteq Z_{\mathbf{VaR}} = \left\{x : \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{VaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta})\right] \leq 0\right\}$$

where $\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right], 0\right\}$.

Outer Approximation

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \geq \mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right) := \min\left\{s : F_{\tilde{X}}(s) \geq 1-\epsilon\right\}.$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right)$.

Theorem (Outer Approximation)

$$Z \subseteq Z_{\mathbf{VaR}} = \left\{x : \mathbb{P}_{\tilde{\zeta}}\left\{\tilde{A}^{\zeta} x \geq \tilde{b}^{\zeta} + \frac{\delta}{\epsilon}\|(x,-1)\|_* \mathbf{e}\right\} \geq 1-\epsilon\right\}$$

Outer Approximation

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \geq \mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right) := \min\left\{s : F_{\tilde{X}}(s) \geq 1-\epsilon\right\}.$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\mathbf{VaR}_{1-\epsilon}\left(\tilde{X}\right)$.

Theorem (Outer Approximation)

$$Z \subseteq Z_{\mathbf{VaR}} = \left\{x : \mathbb{P}_{\tilde{\zeta}}\left\{\tilde{A}^{\zeta}x \geq \tilde{b}^{\zeta} + \frac{\delta}{\epsilon}\|(x,-1)\|_{*}\mathbf{e}\right\} \geq 1-\epsilon\right\}$$

Remarks.

- Asymptotically optimal, i.e., $Z_{\mathbf{VaR}} \to Z$ as $\delta \to 0_+$
- Regular CCP: many existing methods

Inner Approximation: Scenario Approach

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \leq \mathbf{CVaR}_{1}\left(\tilde{X}\right) := \operatorname{ess.sup}(\tilde{X}).$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\operatorname{ess.sup}(\tilde{X})$.

Theorem (Inner Approximation)

$$Z = \left\{x : \frac{\delta}{\epsilon}\|(x,-1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta})\right] \leq 0\right\}$$

where $\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right], 0\right\}$.

Inner Approximation: Scenario Approach

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \leq \mathbf{CVaR}_1\left(\tilde{X}\right) := \operatorname{ess.sup}(\tilde{X}).$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\operatorname{ess.sup}(\tilde{X})$.

Theorem (Inner Approximation)

$$Z \supseteq Z_S = \left\{x : \frac{\delta}{\epsilon}\|(x,-1)\|_* + \operatorname{ess.sup}\left[-\widehat{f}(x,\tilde{\zeta})\right] \leq 0\right\}$$

where $\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right], 0\right\}$.

Inner Approximation: Scenario Approach

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \leq \mathbf{CVaR}_{1}\left(\tilde{X}\right) := \operatorname{ess.sup}(\tilde{X}).$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\operatorname{ess.sup}(\tilde{X})$.

Theorem (Inner Approximation)

$$Z \supseteq Z_S = \left\{\boldsymbol{x} \in \mathbb{R}^n : \mathbb{P}_{\tilde{\zeta}}\left\{\tilde{A}^{\zeta}x \geq \tilde{b}^{\zeta} + \frac{\delta}{\epsilon}\|(x,-1)\|_{*}\mathbf{e}\right\} = 1\right\}$$

Inner Approximation: Scenario Approach

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \leq \mathbf{CVaR}_{1}\left(\tilde{X}\right) := \text{ess.}\sup(\tilde{X}).$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\text{ess.}\sup(\tilde{X})$.

Suppose $\tilde{\zeta}$ has finite support $\{(A^\zeta, b^\zeta)\}_{\zeta\in[N]}$.

Theorem (Inner Approximation)

$$Z \supseteq Z_S = \left\{ \boldsymbol{x} \in \mathbb{R}^n : A^\zeta x \geq b^\zeta + \frac{\delta}{\epsilon}\|(x,-1)\|_* \mathbf{e}, \forall \zeta \in [N] \right\}$$

Remarks.

- Z_S is a conic set
- $Z_S \equiv$ the robust scenario approach (Calafiore and Campi, 2006) to regular CCP when the sample size is small

Inner Approximation: Scenario Approach

Note

$$\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right) \leq \mathbf{CVaR}_{1}\left(\tilde{X}\right) := \text{ess.}\sup(\tilde{X}).$$

Replace $\mathbf{CVaR}_{1-\epsilon}\left(\tilde{X}\right)$ by $\text{ess.}\sup(\tilde{X})$.

Suppose $\tilde{\zeta}$ has finite support $\{(A^{\zeta}, b^{\zeta})\}_{\zeta\in[N]}$.

Theorem (Inner Approximation)

$$Z \supseteq Z_S = \left\{\boldsymbol{x} \in \mathbb{R}^n : A^{\zeta}x \geq b^{\zeta} + \frac{\delta}{\epsilon}\|(x,-1)\|_*\mathbf{e}, \forall \zeta \in [N]\right\}$$

Remarks.

- Z_S is a conic set
- $Z_S \equiv$ the robust scenario approach (Calafiore and Campi, 2006) to regular CCP when the sample size is small
- Z_S can be improved by other less conservative approximations

Inner Approximation: Worst-case CVaR

Note

$$\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^\zeta)^\top x - b_i^\zeta\right], 0\right\} \geq \min_{i\in[m]}\left[(a_i^\zeta)^\top x - b_i^\zeta\right] := \widehat{g}(x,\zeta),$$

Replace $\widehat{f}(x,\zeta)$ by $\widehat{g}(x,\zeta)$ and by monotonicity of coherent risk measure.

Theorem (Inner Approximation)

$$Z = \left\{x : \frac{\delta}{\epsilon}\|(x,-1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta})\right] \leq 0\right\}$$

where $\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^\zeta)^\top x - b_i^\zeta\right], 0\right\}$.

Inner Approximation: Worst-case **CVaR**

Note

$$\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^\zeta)^\top x - b_i^\zeta\right], 0\right\} \geq \min_{i\in[m]}\left[(a_i^\zeta)^\top x - b_i^\zeta\right] := \widehat{g}(x,\zeta),$$

Replace $\widehat{f}(x,\zeta)$ by $\widehat{g}(x,\zeta)$ and by monotonicity of coherent risk measure.

Theorem (Inner Approximation)

$$Z \supseteq Z_C = \left\{x : \frac{\delta}{\epsilon}\|(x,-1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{g}(x,\tilde{\zeta})\right] \leq 0\right\}$$

where $\widehat{g}(x,\zeta) := \min_{i\in[m]}\left[(a_i^\zeta)^\top x - b_i^\zeta\right]$.

Inner Approximation: Worst-case CVaR

Note

$$\widehat{f}(x,\zeta) := \max\left\{\min_{i\in[m]}\left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right], 0\right\} \geq \min_{i\in[m]}\left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right] := \widehat{g}(x,\zeta),$$

Replace $\widehat{f}(x,\zeta)$ by $\widehat{g}(x,\zeta)$ and by monotonicity of coherent risk measure.

Theorem (Inner Approximation)

$$Z \supseteq Z_C = \left\{x : \frac{\delta}{\epsilon}\|(x,-1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{g}(x,\tilde{\zeta})\right] \leq 0\right\}$$

where $\widehat{g}(x,\zeta) := \min_{i\in[m]}\left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right]$.

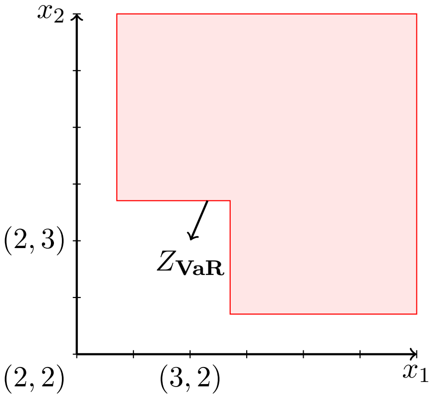
Remarks.

- Z_C is a conic set
- $Z_C \equiv$ the worst-case **CVaR** approximation of DRCCP-W (Nemirovski and Shapiro, 2006)

Illustrations of $Z, Z_{\mathbf{VaR}}, Z_C, Z_S$

Theorem (Model Comparison)

$$Z_S \subseteq Z_C \subseteq Z \subseteq Z_{\mathbf{VaR}}.$$



Consider

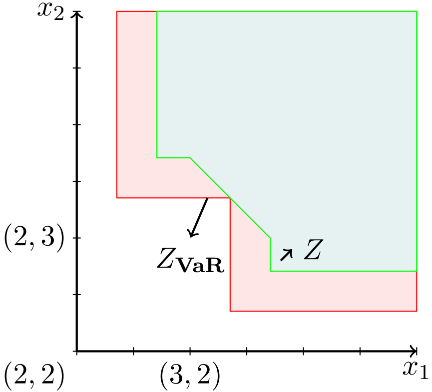
$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P} \left\{ \begin{array}{l} \tilde{a}_1 \le x_1 \\ \tilde{a}_2 \le x_2 \end{array} \right\} \ge 1 - \epsilon \right\}$$

- Risk parameter $\epsilon = 2/3$
- Wasserstein radius $\delta = 1/6$
- $N = 3$ empirical data points:
  $(a_1^1, a_2^1) = (1, 3)$
  $(a_1^2, a_2^2) = (3, 1)$
  $(a_1^3, a_2^3) = (2, 2)$

Illustrations of $Z, Z_{\textbf{VaR}}, Z_C, Z_S$

Theorem (Model Comparison)

$$Z_S \subseteq Z_C \subseteq Z \subseteq Z_{\textbf{VaR}}.$$



Consider

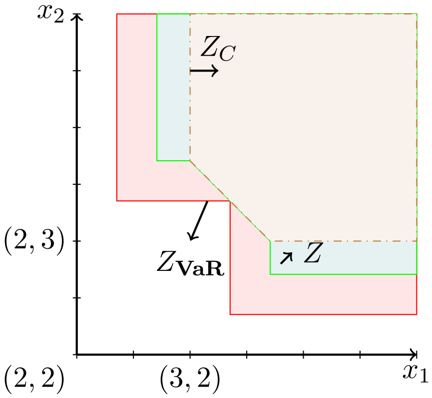
$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P} \left\{ \begin{array}{c} \tilde{a}_1 \leq x_1 \\ \tilde{a}_2 \leq x_2 \end{array} \right\} \geq 1 - \epsilon \right\}$$

- Risk parameter $\epsilon = 2/3$
- Wasserstein radius $\delta = 1/6$
- $N = 3$ empirical data points:
  $(a_1^1, a_2^1) = (1, 3)$
  $(a_1^2, a_2^2) = (3, 1)$
  $(a_1^3, a_2^3) = (2, 2)$

Illustrations of $Z, Z_{\mathbf{VaR}}, Z_C, Z_S$

Theorem (Model Comparison)

$$Z_S \subseteq Z_C \subseteq Z \subseteq Z_{\mathbf{VaR}}.$$



Consider

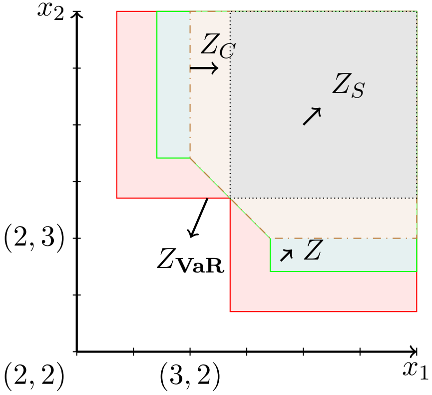
$$Z = \left\{ x : \inf_{\mathbb{P}\in\mathcal{P}^W} \mathbb{P}\left\{ \begin{array}{c} \tilde{a}_1 \le x_1 \\ \tilde{a}_2 \le x_2 \end{array} \right\} \ge 1-\epsilon \right\}$$

- Risk parameter $\epsilon = 2/3$
- Wasserstein radius $\delta = 1/6$
- $N = 3$ empirical data points:
  $(a_1^1, a_2^1) = (1,3)$
  $(a_1^2, a_2^2) = (3,1)$
  $(a_1^3, a_2^3) = (2,2)$

Illustrations of $Z, Z_{\mathbf{VaR}}, Z_C, Z_S$

Theorem (Model Comparison)

$$Z_S \subseteq Z_C \subseteq Z \subseteq Z_{\mathbf{VaR}}.$$



Consider

$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P} \left\{ \begin{array}{c} \tilde{a}_1 \leq x_1 \\ \tilde{a}_2 \leq x_2 \end{array} \right\} \geq 1 - \epsilon \right\}$$

- Risk parameter $\epsilon = 2/3$
- Wasserstein radius $\delta = 1/6$
- $N = 3$ empirical data points:
  $(a_1^1, a_2^1) = (1, 3)$
  $(a_1^2, a_2^2) = (3, 1)$
  $(a_1^3, a_2^3) = (2, 2)$

Mixed Integer Program Reformulation

CVaR Reformulation: Recall

DRCCP-W set

$$Z = \left\{ x : \inf_{\mathbb{P} \in \mathcal{P}^W} \mathbb{P}\left\{ \tilde{A}x \geq \tilde{b} \right\} \geq 1 - \epsilon \right\}.$$

Theorem (Exact Formulation)

$$Z = \left\{ x : \frac{\delta}{\epsilon} \|(x, 1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

where $\widehat{f}(x, \zeta) := \max\left\{ \min_{i \in [m]} \left[(a_i^{\zeta})^\top x - b_i^{\zeta} \right], 0 \right\}$.

Mixed Integer Program (MIP) Reformulation

DRCCP-W set

$$Z = \left\{ x : \frac{\delta}{\epsilon} \|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon} \left[-\widehat{f}(x,\tilde{\zeta}) \right] \leq 0 \right\},$$

where $\widehat{f}(x,\zeta) := \max \left\{ \min_{i\in[m]} \left[(a_i^\zeta)^\top x - b_i^\zeta \right], 0 \right\}.$

Mixed Integer Program (MIP) Reformulation

DRCCP-W set

$$Z = \left\{ x : \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta})\right] \leq 0 \right\},$$

where $\widehat{f}(x,\boldsymbol{\zeta}) := \max\left\{ \min_{i\in[m]} \left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right], 0 \right\}$.

- Linearize outer maximum with binary variable $z^\zeta \in \{0,1\}$ and continuous variable w^ζ as $\widehat{f}(x,\boldsymbol{\zeta}) = w^\zeta$ and

$$w^\zeta = \begin{cases} 0, & \text{if } \min_{i\in[m]} \left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right] \leq 0 \\ \min_{i\in[m]} \left[(a_i^{\zeta})^\top x - b_i^{\zeta}\right], & \text{otherwise} \end{cases}$$

Mixed Integer Program (MIP) Reformulation

DRCCP-W set

$$Z = \left\{ x : \frac{\delta}{\epsilon} \|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x, \tilde{\zeta}) \right] \leq 0 \right\},$$

where $\widehat{f}(x, \boldsymbol{\zeta}) := \max \left\{ \min_{i \in [m]} \left[(a_i^\zeta)^\top x - b_i^\zeta \right], 0 \right\}.$

- Linearize outer maximum with binary variable $z^\zeta \in \{0,1\}$ and continuous variable w^ζ as $\widehat{f}(x, \boldsymbol{\zeta}) = w^\zeta$ and

$$0 \leq w^\zeta \leq M^\zeta z^\zeta$$
$$w^\zeta - M^\zeta (1 - z^\zeta) \leq (a_i^\zeta)^\top x - b_i^\zeta, \forall i \in [m]$$

where $M^\zeta \geq \max_{x \in Z} \min_{i \in [m]} \left[\left| (a_i^\zeta)^\top x - b_i^\zeta \right| \right]$

Mixed Integer Program (MIP) Reformulation

DRCCP-W set

$$Z = \left\{ x : \begin{array}{l} \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-w^\zeta\right] \leq 0 \\ 0 \leq w^\zeta \leq M^\zeta z^\zeta, \forall \zeta \\ w^\zeta - M^\zeta(1-z^\zeta) \leq (a_i^\zeta)^\top x - b_i^\zeta, \forall i \in [m], \forall \zeta \\ z^\zeta \in \{0,1\}, \forall \zeta \end{array} \right\}$$

- Empirical distribution is finite-support $\{(A^\zeta, b^\zeta)\}_{\zeta \in [N]} \Rightarrow$ set Z is an MIP
- Optimality is guaranteed by the solvers

Mixed Integer Program (MIP) Reformulation

DRCCP-W set

$$Z = \left\{ x : \begin{array}{l} \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-w^\zeta\right] \leq 0 \\ 0 \leq w^\zeta \leq M^\zeta z^\zeta, \forall \zeta \\ w^\zeta - M^\zeta(1 - z^\zeta) \leq (a_i^\zeta)^\top x - b_i^\zeta, \forall i \in [m], \forall \zeta \\ z^\zeta \in \{0,1\}, \forall \zeta \end{array} \right\}$$

- Empirical distribution is finite-support $\{(A^\zeta, b^\zeta)\}_{\zeta \in [N]} \Rightarrow$ set Z is an MIP
- Optimality is guaranteed by the solvers
- Similar to regular CCP, (1) big M coefficients weaken the formulation, (2) number of binary variables grows as sample size N increases
  - Both will be addressed for binary DRCCP

Binary DRCCP-W and Submodularity

Preliminaries

Binary DRCCP-W set

$$Z = \left\{ x \in \{0,1\}^n : \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta}) \right] \leq 0 \right\},$$

where $\widehat{f}(x,\zeta) := \min_{i\in[m]} \left\{ \max\left[(a_i^{\zeta})^\top x - b_i^{\zeta}, 0 \right] \right\}.$

Preliminaries

Binary DRCCP-W set

$$Z = \left\{ x \in \{0,1\}^n : \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta})\right] \le 0 \right\},$$

where $\widehat{f}(x,\boldsymbol{\zeta}) := \min_{i\in[m]} \left\{ \max\left[(a_i^{\zeta})^\top x - b_i^{\zeta}, 0 \right] \right\}$.

Fact 1

Given $\boldsymbol{d}_1 \in \mathbb{R}_+^n, d_2, d_3 \in \mathbb{R}$, function $f(x) = -\max\left(\boldsymbol{d}_1^\top x + d_2, d_3\right)$ is submodular over the binary hypercube.

Fact 2 (Edmonds, 1970)

For a submodular function $f : \{0,1\}^n \to \mathbb{R}$,
$\text{conv}(\text{epi}(f)) = \text{conv}\left\{(x,w) : f(x) \le w, x \in \{0,1\}^n\right\} =$ "extended polymatroid inequalities" (EPI)

The time complexity of separation over EPI is $O(n\log(n))$

Binary DRCCP-W: Submodular Constrained Reformulation

Binary DRCCP-W set

$$Z = \left\{ x \in \{0,1\}^n : \frac{\delta}{\epsilon} \|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[-\widehat{f}(x,\tilde{\zeta}) \right] \leq 0 \right\},$$

where $\widehat{f}(x,\zeta) := \min_{i\in[m]} \left\{ \max\left[(a_i^{\zeta})^\top x - b_i^{\zeta}, 0 \right] \right\}.$

Binary DRCCP-W: Submodular Constrained Reformulation

Binary DRCCP-W set

$$Z = \left\{ x \in \{0,1\}^n : \begin{array}{l} \dfrac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[w^{\tilde{\xi}}\right] \le 0 \\ -\max\left[(a_i^\zeta)^\top x - b_i^\zeta, 0\right] \le w^\zeta, \forall i \in [m], \forall \zeta \end{array} \right\}$$

- Let $w^\zeta = \widehat{f}(x, \boldsymbol{\zeta})$ and linearize it

Binary DRCCP-W: Submodular Constrained Reformulation

Binary DRCCP-W set

$$Z = \left\{ x: \begin{array}{l} \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[w^{\tilde{\xi}}\right] \leq 0 \\ \boxed{-\max\left[(\widehat{a}_i^{\zeta,x})^\top x + (\widehat{a}_i^{\zeta,y})^\top y - \widehat{b}_i^{\zeta}, 0\right] \leq w^{\zeta}}, \forall i \in [m], \forall \zeta \\ x_r + y_r = 1, \forall r \in [n], \\ x, y \in \{0,1\}^n \end{array} \right\}$$

- Let $w^{\zeta} = \widehat{f}(x, \boldsymbol{\zeta})$ and linearize it
- Let $y_r = 1 - x_r$ and choose vectors $\widehat{a}_i^{\zeta,x}, \widehat{a}_i^{\zeta,y} \in \mathbb{R}_+^n$ such that

$$(a_i^{\zeta})^\top x - b_i^{\zeta} = (\widehat{a}_i^{\zeta,x})^\top x + (\widehat{a}_i^{\zeta,y})^\top y - \widehat{b}_i^{\zeta}$$

Binary DRCCP-W: Submodular Constrained Reformulation

Binary DRCCP-W set

$$Z = \left\{ x : \begin{array}{l} \frac{\delta}{\epsilon}\|(x,1)\|_* + \mathbf{CVaR}_{1-\epsilon}\left[w^{\tilde{\zeta}}\right] \le 0 \\ \boxed{-\max\left[(\widehat{a}_i^{\zeta,x})^\top x + (\widehat{a}_i^{\zeta,y})^\top y - \widehat{b}_i^\zeta, 0\right] \le w^\zeta}, \forall i \in [m], \forall \zeta \\ x_r + y_r = 1, \forall r \in [n], \\ x, y \in \{0,1\}^n \end{array} \right\}$$

- Let $w^\zeta = \widehat{f}(x, \boldsymbol{\zeta})$ and linearize it
- Let $y_r = 1 - x_r$ and choose vectors $\widehat{a}_i^{\zeta,x}, \widehat{a}_i^{\zeta,y} \in \mathbb{R}_+^n$ such that

$$(a_i^\zeta)^\top x - b_i^\zeta = (\widehat{a}_i^{\zeta,x})^\top x + (\widehat{a}_i^{\zeta,y})^\top y - \widehat{b}_i^\zeta$$

- Facts 1 and 2⇒(1) Z is submodular constrained set and (2)separation of these constraints is very efficient

Numerical Illustration : Setting

Consider distributionally robust chance constrained knapsack problem

$$
\begin{aligned}
v^* = \max_{\boldsymbol{x}} \quad & \boldsymbol{c}^\top \boldsymbol{x}, \\
\text{s.t.} \quad & \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\left\{ \tilde{a}_i^\top x \le \tilde{b}^i, \forall i \in [m] \right\} \ge 1 - \epsilon.
\end{aligned}
$$

- Let $n = 20, m = 10$
- Generate 10 random instances and for each instance, there are $N = 100$ samples.

Results (1): Continuous Knapsack $x \in [0,1]^n$

ϵ	δ	Instances	BigM Model		VaR Model			CVaR Model		
			Opt.Val	Time	Value	GAP	Time	Value	GAP	Time
0.05	0.01	1	54.93	6.11	56.37	2.62%	3.37	54.30	1.14%	0.06
		2	47.69	5.24	48.79	2.29%	2.04	47.16	1.11%	0.05
		3	50.73	4.44	51.43	1.38%	4.43	50.38	0.70%	0.05
		4	53.97	3.61	54.98	1.87%	4.75	52.72	2.32%	0.06
		5	54.96	6.99	56.44	2.68%	4.20	52.88	3.79%	0.05
		6	56.03	6.46	57.40	2.44%	2.64	54.97	1.89%	0.05
		7	54.17	6.69	55.04	1.62%	3.68	53.26	1.67%	0.05
		8	55.40	5.81	56.55	2.09%	3.19	54.15	2.26%	0.05
		9	57.63	4.91	58.95	2.29%	4.20	57.07	0.96%	0.05
		10	56.31	4.34	57.15	1.50%	4.71	55.95	0.63%	0.06
Average				5.46		2.08%	3.72		1.65%	0.05
0.05	0.02	1	53.97	3.94	55.92	3.63%	3.27	53.83	0.24%	0.05
		2	47.05	3.63	48.42	2.92%	3.20	46.79	0.53%	0.04
		3	50.12	5.26	51.02	1.79%	4.48	49.96	0.33%	0.05
		4	52.98	5.14	54.49	2.84%	4.83	52.28	1.33%	0.06
		5	54.10	3.76	55.95	3.41%	3.67	52.44	3.07%	0.05
		6	55.16	6.02	56.90	3.16%	3.33	54.52	1.17%	0.05
		7	53.41	3.91	54.55	2.13%	3.81	52.83	1.08%	0.05
		8	54.47	2.77	56.09	2.98%	3.34	53.71	1.39%	0.06
		9	56.85	3.40	58.44	2.79%	4.00	56.59	0.46%	0.05
		10	55.65	5.47	56.71	1.90%	4.90	55.53	0.22%	0.06
Average				4.33		2.76%	3.88		0.98%	0.05

Results (2): Testing Robustness

Instances	DRCCP Model			CCP Model		Target Violation (ϵ)
	δ^*	Opt.Val	90-Percentile Violation	Opt.Val	90-Percentile Violation	
1	0.03	53.76	0.042	56.99	0.135	
2	0.02	50.06	0.044	52.67	0.087	
3	0.03	52.37	0.031	55.11	0.153	
4	0.01	56.94	0.039	58.33	0.096	
5	0.02	53.38	0.028	55.89	0.121	
6	0.02	50.25	0.032	52.13	0.096	0.05
7	0.01	59.38	0.047	60.98	0.080	
8	0.03	54.60	0.047	57.77	0.129	
9	0.03	62.51	0.047	66.39	0.118	
10	0.03	52.82	0.036	56.90	0.132	

Results (3): Binary Knapsack $\boldsymbol{x} \in \{0, 1\}^n$

ϵ	δ	Instances	n	I	MIP Formulation				Submodular Formulation	
					UB	LB	Time	GAP	Opt. Val.	Time
0.05	0.1	1	20	10	93	86	3600.0	7.5%	89	49.3
		2	20	10	97	90	3600.0	7.2%	95	30.6
		3	20	10	95	84	3600.0	11.6%	90	387.0
		4	20	10	84	74	3600.0	11.9%	78	275.7
		5	20	10	87	81	3600.0	6.9%	82	140.4
		6	20	10	97	85	3600.0	12.4%	88	972.5
		7	20	10	89	75	3600.0	15.7%	84	169.6
		8	20	10	100	88	3600.0	12.0%	96	80.5
		9	20	10	96	78	3600.0	18.8%	92	59.3
		10	20	10	93	93	3542.7	0.0%	93	18.2
Average							3594.3	10.4%		218.3
0.1	0.1	1	20	10	100	NA	3600.0	NA	92	172.9
		2	20	10	106	NA	3600.0	NA	99	164.0
		3	20	10	105	87	3600.0	17.1%	93	569.1
		4	20	10	92	67	3600.0	27.2%	82	600.5
		5	20	10	95	NA	3600.0	NA	86	332.0
		6	20	10	109	NA	3600.0	NA	94	1852.4
		7	20	10	96	NA	3600.0	NA	88	279.8
		8	20	10	108	82	3600.0	24.1%	100	133.2
		9	20	10	102	NA	3600.0	NA	94	389.3
		10	20	10	103	96	3600.0	6.8%	96	149.7
Average							3600.0	18.8%		464.3

Concluding Remarks

Concluding Remarks

- DRCCP-W admits a **CVaR** interpretation
  - Derive inner and outer approximations
- DRCCP-W is mixed integer program representable
  - With big-M coefficients and additional binary variables
- Binary DRCCP-W $\equiv$ a submodular constrained optimization problem
  - Without big-M coefficients or additional binary variables

References:

- **W. Xie**. "On Distributionally Robust Chance Constrained Program with Wasserstein Distance". Available at Optimization Online, 2018.

Thank you!