Relaxations and approximations of chance constraints under finite distributions

Shabbir Ahmed, Weijun Xie

Friday Seminar

Li Chen

November 19, 2021

Chance-Constrained Program

Consider a chance-constrained Program (CCP):

$$v^* = \min_{\boldsymbol{x} \in \mathbb{R}^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \mathbb{P}\left[\tilde{\boldsymbol{\xi}} : \boldsymbol{x} \in \mathcal{X}(\tilde{\boldsymbol{\xi}}) \right] \geq 1 - \epsilon \right\} \tag{1}$$

where the random vector $\tilde{\boldsymbol{\xi}}$ has support set Ξ.

Chance-Constrained Program

Consider a chance-constrained Program (CCP):

$$v^* = \min_{\boldsymbol{x} \in \mathbb{R}^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \mathbb{P}\left[\tilde{\boldsymbol{\xi}} : \boldsymbol{x} \in \mathcal{X}(\tilde{\boldsymbol{\xi}}) \right] \geq 1 - \epsilon \right\} \tag{1}$$

where the random vector $\tilde{\boldsymbol{\xi}}$ has support set Ξ.

Assumptions

- Finite scenarios: $\mathbb{P}\left[\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}^i\right] = p_i$ for $i \in [N]$.
- The feasible region in scenario i has the form of

  $$\mathcal{X}(\boldsymbol{\xi}^i) = \left\{ \boldsymbol{x} : \boldsymbol{A}^i \boldsymbol{x} \geq \boldsymbol{b}^i \right\}$$

  where $\boldsymbol{\xi}^i \triangleq (\boldsymbol{A}^i, \boldsymbol{b}^i)$.
- The sets $\mathcal{X}(\boldsymbol{\xi}^i)$ is non-empty and compact (can be relaxed to sharing the same recession cone)

Challenges

- Typically, CCP involves two challenges:
  - Check feasibility needs computing probability (not an issue here)
  - Optimization over non-convex feasible region

[1] James Luedtke, Shabbir Ahmed, and George L Nemhauser. "An integer programming approach for linear programs with probabilistic constraints". In: *Mathematical programming* 122.2 (2010), pp. 247–272.

Challenges

- Typically, CCP involves two challenges:
  - Check feasibility needs computing probability (not an issue here)
  - Optimization over non-convex feasible region
- The chance-constrained program (1) is strongly NP-hard.[1]
- Mixed integer linear programming (MILP) formulation:

$$v^* = \min_{\boldsymbol{x},\boldsymbol{z}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{ll} \boldsymbol{A}^i \boldsymbol{x} \geq z_i \boldsymbol{b}^i - (1 - z_i)\boldsymbol{M}_i & \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon & \\ \boldsymbol{z} \in \{0,1\}^N & \end{array} \right\} \tag{2}$$

[1] Luedtke, Ahmed, and Nemhauser, "An integer programming approach for linear programs with probabilistic constraints".

Challenges

- Typically, CCP involves two challenges:
  - Check feasibility needs computing probability (not an issue here)
  - Optimization over non-convex feasible region
- The chance-constrained program (1) is strongly NP-hard.[1]
- Mixed integer linear programming (MILP) formulation:

$$v^* = \min_{\boldsymbol{x},\boldsymbol{z}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \boldsymbol{A}^i \boldsymbol{x} \geq z_i \boldsymbol{b}^i - (1 - z_i)\boldsymbol{M}_i \quad \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\} \tag{2}$$

  - The above big-M formulation is difficult to solve as its continuous relaxation is very weak

[1] Luedtke, Ahmed, and Nemhauser, "An integer programming approach for linear programs with probabilistic constraints".

Solution Approaches

1. Improving the MILP relaxation
    - Coefficient tightening
    - Quantile bound
    - Nonanticipative relaxations
    - Quantile cut and quantile closure
2. Constructing convex approximation
    - CVaR approximation
    - Scenario approximation
    - Bisection heuristic
    - Bicriteria approximation

Coefficient Tightening

Idea: estimate the big-M coefficients and reduce them recursively

- Each row of the constraint corresponding to scenario i is

$$\boldsymbol{a}_\ell^{i\top}\boldsymbol{x} \geq b_\ell^i z_i - M_\ell^i(1 - z_i)$$

where the big-M coefficient

$$-M_\ell^i \leq \inf\left\{\boldsymbol{a}_\ell^{i\top}\boldsymbol{x} : (\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{X}\right\}$$

and $\mathcal{X}$ is the feasible region in MILP formulation (2).

Coefficient Tightening

Idea: estimate the big-M coefficients and reduce them recursively

- Each row of the constraint corresponding to scenario i is

$$\boldsymbol{a}_\ell^{i\top}\boldsymbol{x} \geq b_\ell^i z_i - M_\ell^i(1 - z_i)$$

  where the big-M coefficient

$$-M_\ell^i \leq \inf\left\{\boldsymbol{a}_\ell^{i\top}\boldsymbol{x} : (\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{X}\right\}$$

  and $\mathcal{X}$ is the feasible region in MILP formulation (2).
- Obtaining the optimal value itself is hard $\Longrightarrow$ using the continuous relaxation of $\mathcal{X}$
- Start from very large M_ℓ^i, and replace it by the optimal value of the relaxed problem, and proceed recursively

Quantile Bound

Assume $\mathcal{X}(\boldsymbol{\xi}^i) \subseteq S$ for all $i \in [N]$ where S is compact.

Quantile Bound

Assume $\mathcal{X}(\boldsymbol{\xi}^i) \subseteq S$ for all $i \in [N]$ where S is compact.

For simplicity, assume $p_i = 1/N$ for now, then the CCP becomes

$$\min_{\boldsymbol{x}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \sum_{i \in [N]} \mathbb{I}(\boldsymbol{x} \in \mathcal{X}(\boldsymbol{\xi}^i)) \geq (1-\epsilon)N \right\}$$

Let $k := \lfloor \epsilon N \rfloor$,

- Compute
$$\beta^i = \min \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in S \cap \mathcal{X}(\boldsymbol{\xi}^i) \right\}$$
- Sort $\beta^1 \geq \beta^2 \geq ... \geq \beta^N$
- Then we must have
$$v^* \geq \beta^{k+1}$$
because a feasible $\boldsymbol{x}$ must satisfy $\sum_{i \in [N]} \mathbb{I}(\boldsymbol{x} \in \mathcal{X}(\boldsymbol{\xi}^i)) \geq N - k$, which implies $\boldsymbol{x} \in S \cap \mathcal{X}(\boldsymbol{\xi}^i)$ for some $i \in \{1, 2, ..., k+1\}$.

Extend to general distribution p_i

Nonanticipative Relaxations

Idea: variable splitting

$$v^* = \min_{\boldsymbol{x}^1,\ldots,\boldsymbol{x}^N,\boldsymbol{z}} \left\{ \sum_{i\in[N]} p_i \boldsymbol{c}^\top \boldsymbol{x}^i : \begin{array}{l} \sum_{i\in[N]} \boldsymbol{H}^i \boldsymbol{x}^i = \boldsymbol{h} \\ \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1 - z_i)\boldsymbol{M}^i \quad \forall i \in [N] \\ \sum_{i\in[N]} p_i z_i \geq 1 - \epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}$$

where

$$\sum_{i\in[N]} \boldsymbol{H}^i \boldsymbol{x}^i = \boldsymbol{h} \iff \boldsymbol{x}^1 = \ldots = \boldsymbol{x}^N$$

is the system of nonanticipativity constraints.

Idea: variable splitting

$$v^* = \min_{\boldsymbol{x}^1,\ldots,\boldsymbol{x}^N,\boldsymbol{z}} \left\{ \sum_{i\in[N]} p_i \boldsymbol{c}^\top \boldsymbol{x}^i : \begin{array}{l} \sum_{i\in[N]} \boldsymbol{H}^i \boldsymbol{x}^i = \boldsymbol{h} \\ \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1-z_i)\boldsymbol{M}^i \quad \forall i \in [N] \\ \sum_{i\in[N]} p_i z_i \geq 1-\epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}$$

where

$$\sum_{i\in[N]} \boldsymbol{H}^i \boldsymbol{x}^i = \boldsymbol{h} \iff \boldsymbol{x}^1 = \ldots = \boldsymbol{x}^N$$

is the system of nonanticipativity constraints.

- Lagrangian relaxation: A lower bound of v^* is

$$v_1^{LD} = \sup_{\boldsymbol{\lambda}} \mathcal{L}_1(\boldsymbol{\lambda}) - \boldsymbol{\lambda}^\top \boldsymbol{h}$$

where

$$\mathcal{L}_1(\boldsymbol{\lambda}) = \min_{\boldsymbol{x},\boldsymbol{z}} \left\{ \sum_{i\in[N]} p_i \boldsymbol{c}^\top \boldsymbol{x}^i + \boldsymbol{\lambda}^\top \boldsymbol{H}^i \boldsymbol{x}^i : \begin{array}{l} \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1-z_i)\boldsymbol{M}^i \\ \sum_{i\in[N]} p_i z_i \geq 1-\epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}$$

Apply subgradient to optimize $\boldsymbol{\lambda}$.

Nonanticipative Relaxations

Note also

$$v^* = \min_{y,\boldsymbol{x},\boldsymbol{z}} \left\{ y : \begin{array}{ll} y \geq \boldsymbol{c}^\top \boldsymbol{x}^i & \forall i \in [N] \\ \sum_{i \in [N]} \boldsymbol{H}^i \boldsymbol{x}^i = \boldsymbol{h} & \\ \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1 - z_i) \boldsymbol{M}^i & \forall i \in [N] \\ \sum_{i \in [N]} p_i z_i \geq 1 - \epsilon & \\ \boldsymbol{z} \in \{0,1\}^N & \end{array} \right\} \tag{3}$$

Nonanticipative Relaxations

Note also

$$v^* = \min_{y,\boldsymbol{x},\boldsymbol{z}} \left\{ y : \begin{array}{ll} y \geq \boldsymbol{c}^\top \boldsymbol{x}^i & \forall i \in [N] \\ \sum_{i\in[N]} \boldsymbol{H}^i \boldsymbol{x}^i = \boldsymbol{h} & \\ \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1 - z_i)\boldsymbol{M}^i & \forall i \in [N] \\ \sum_{i\in[N]} p_i z_i \geq 1 - \epsilon & \\ \boldsymbol{z} \in \{0,1\}^N & \end{array} \right\} \tag{3}$$

- Lagrangian relaxation: Another lower bound of v^* is

$$v_2^{LD} = \min_y \sup_{\boldsymbol{\lambda}} y + \mathcal{L}_2(\boldsymbol{\lambda}) - \boldsymbol{\lambda}^\top \boldsymbol{h}$$

where

$$\mathcal{L}_2(\boldsymbol{\lambda}) = \min_{\boldsymbol{x},\boldsymbol{z}} \left\{ \boldsymbol{\lambda}^\top \boldsymbol{H}^i \boldsymbol{x}^i : \begin{array}{l} y \geq \boldsymbol{c}^\top \boldsymbol{x}^i \\ \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1 - z_i)\boldsymbol{M}^i \\ \sum_{i\in[N]} p_i z_i \geq 1 - \epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}$$

Bisection on y and apply subgradient to optimize $\boldsymbol{\lambda}$.

Relation between The Relaxations

Let v^M be the optimal value of LP relaxation of the MILP (2) , then

$$v^M \leq v_1^{LD} \leq v_2^{LD}$$

Relation between The Relaxations

Let v^M be the optimal value of LP relaxation of the MILP (2) , then

$$v^M \leq v_1^{LD} \leq v_2^{LD}$$

Proof.

$$v^M = \sup_{\boldsymbol{\lambda}} \min_{y,\boldsymbol{x}^i,\boldsymbol{z}} \left\{ y + \boldsymbol{\lambda}^\top \boldsymbol{H}^i \boldsymbol{x}^i : \begin{array}{l} y \geq \sum_{i\in[N]} p_i \boldsymbol{c}^\top \boldsymbol{x}^i \\ \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1-z_i)\boldsymbol{M}^i \\ \sum_{i\in[N]} p_i z_i \geq 1-\epsilon \\ \boldsymbol{z} \in [0,1]^N \end{array} \right\} - \boldsymbol{\lambda}^\top \boldsymbol{h}$$

$$v_1^{LD} = \sup_{\boldsymbol{\lambda}} \min_{y,\boldsymbol{x}^i,\boldsymbol{z}} \left\{ y + \boldsymbol{\lambda}^\top \boldsymbol{H}^i \boldsymbol{x}^i : \begin{array}{l} y \geq \sum_{i\in[N]} p_i \boldsymbol{c}^\top \boldsymbol{x}^i \\ \boldsymbol{A}^i \boldsymbol{x}^i \geq z_i \boldsymbol{b}^i - (1-z_i)\boldsymbol{M}^i \\ \sum_{i\in[N]} p_i z_i \geq 1-\epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\} - \boldsymbol{\lambda}^\top \boldsymbol{h}$$

Relation between The Relaxations

Let v^M be the optimal value of LP relaxation of the MILP (2) , then

$$v^M \le v_1^{LD} \le v_2^{LD}$$

Proof.

$$v_1^{LD} = \sup_{\boldsymbol{\lambda}} \min_{y, \boldsymbol{x}^i, \boldsymbol{z}} \left\{ y + \boldsymbol{\lambda}^\top \boldsymbol{H}^i \boldsymbol{x}^i : \begin{array}{l} y \ge \sum_{i\in[N]} p_i \boldsymbol{c}^\top \boldsymbol{x}^i \\ \boldsymbol{A}^i \boldsymbol{x}^i \ge z_i \boldsymbol{b}^i - (1 - z_i)\boldsymbol{M}^i \\ \sum_{i\in[N]} p_i z_i \ge 1 - \epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\} - \boldsymbol{\lambda}^\top \boldsymbol{h}$$

$$v_2^{LD} = \min_{y} \sup_{\boldsymbol{\lambda}} \min_{\boldsymbol{x}^i, \boldsymbol{z}} \left\{ y + \boldsymbol{\lambda}^\top \boldsymbol{H}^i \boldsymbol{x}^i : \begin{array}{l} y \ge \max_{i\in[N]} \boldsymbol{c}^\top \boldsymbol{x}^i \\ \boldsymbol{A}^i \boldsymbol{x}^i \ge z_i \boldsymbol{b}^i - (1 - z_i)\boldsymbol{M}^i \\ \sum_{i\in[N]} p_i z_i \ge 1 - \epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\} - \boldsymbol{\lambda}^\top \boldsymbol{h}$$

Comparison of The Relaxations

$\mathcal{X}(\boldsymbol{\xi}^i) = \{\boldsymbol{x} \in [0,1]^n : \boldsymbol{A}^i \boldsymbol{x} \leq \boldsymbol{b}^i\}$

Table 1: Bound comparison for multi-dimensional continuous knapsack instances.

			No big-M Str.			With big-M Str.			v^Q	v^H
Instance	ϵ	N	$v^C(M)$	v_1^{LD}	v_2^{LD}	$v^C(M)$	v_1^{LD}	v_2^{LD}		
mk-20-10	0.1	100	10.1%	7.3%	1.3%	2.3%	2.3%	1.2%	2.4%	0.4%
		500	10.0%	7.0%	1.4%	2.4%	2.4%	1.2%	2.1%	0.2%
		1000	10.0%	7.3%	1.6%	2.5%	2.5%	1.4%	2.5%	0.3%
		3000	9.8%	7.2%	1.7%	2.6%	2.6%	1.5%	2.5%	0.2%
	0.2	100	14.5%	10.5%	1.3%	3.0%	3.0%	1.1%	2.0%	0.6%
		500	14.7%	10.3%	1.4%	3.0%	2.9%	1.3%	2.1%	0.3%
		1000	14.8%	10.7%	1.7%	3.2%	3.2%	1.5%	2.5%	0.3%
		3000	14.4%	10.5%	1.8%	3.2%	3.2%	1.6%	2.6%	0.1%
mk-39-5	0.1	100	8.0%	7.4%	2.0%	2.0%	2.0%	1.5%	3.2%	0.6%
		500	8.9%	8.3%	2.5%	2.6%	2.5%	2.0%	3.8%	0.3%
		1000	8.8%	8.3%	2.6%	2.6%	2.6%	2.1%	3.9%	0.4%
		3000	8.7%	8.3%	2.9%	2.8%	2.8%	2.3%	4.3%	0.1%
	0.2	100	11.4%	10.7%	2.1%	2.7%	2.7%	1.8%	3.3%	0.4%
		500	12.4%	11.7%	2.6%	3.4%	3.3%	2.2%	3.5%	0.3%
		1000	12.4%	11.7%	2.9%	3.5%	3.4%	2.4%	4.0%	0.2%
		3000	12.1%	11.6%	3.0%	3.5%	3.5%	2.5%	4.2%	0.1%

Comparison of The Relaxations

$$\mathcal{X}(\boldsymbol{\xi}^i) = \{\boldsymbol{x} \in \{0,1\}^n : \boldsymbol{A}^i \boldsymbol{x} \leq \boldsymbol{b}^i\}$$

Table 4: Bound comparison for multi-dimensional binary knapsack instances.

Instances	ϵ	N	$\underline{v}^C(M)$	z_1^{LP}	z_2^{NLP}	v^Q	v^{QG}	v^H
mk-20-10	0.1	100	3.5%	3.5%	2.3%	1.6%	1.1%	0.0%
		500	3.8%	3.8%	2.6%	1.8%	1.5%	0.0%
		1000	3.8%	3.8%	2.7%	2.0%	1.8%	0.0%
		3000	3.8%	3.8%	2.7%	2.0%	1.7%	0.0%
	0.2	100	4.8%	4.7%	2.8%	2.3%	2.4%	0.0%
		500	3.9%	3.9%	2.2%	1.5%	1.6%	0.0%
		1000	4.4%	4.4%	2.7%	2.2%	2.2%	0.0%
		3000	4.4%	4.4%	2.7%	2.1%	2.0%	0.0%
mk-39-5	0.1	100	3.2%	3.2%	2.7%	3.3%	2.2%	0.6%
		500	≤3.9%*	≤3.9%	≤3.3%	≤3.9%	≤3.1%	≤0.3%
		1000	≤4.0%	≤4.0%	≤3.5%	≤4.1%	≤3.4%	≤1.6%
		3000	≤4.0%	≤4.0%	≤3.5%	≤4.2%	≤3.3%	≤2.7%
	0.2	100	3.9%	3.9%	2.9%	3.4%	3.0%	0.4%
		500	≤4.2%	≤4.2%	≤3.0%	≤3.3%	≤3.6%	≤1.1%
		1000	≤4.4%	≤4.3%	≤3.3%	≤3.8%	≤4.0%	≤2.3%
		3000	≤4.5%	≤4.5%	≤3.5%	≤4.0%	≤3.8%	≤3.2%

* A "≤" indicates instances for which the optimal value is not known, and the associated number represents an upper bound on the true optimality gap.

Quantile Cut

- For any $\boldsymbol{\alpha} \in \mathbb{R}^n$, let

$$\beta^i_{\boldsymbol{\alpha}}(S) = \min\left\{\boldsymbol{\alpha}^\top \boldsymbol{x} : \boldsymbol{x} \in S \cap \mathcal{X}(\boldsymbol{\xi}^i)\right\}$$

- Let $\beta^q_{\boldsymbol{\alpha}}(S)$ be the $(1-\epsilon)$-quantile $\beta^i_{\boldsymbol{\alpha}}$, then

$$\boldsymbol{\alpha}^\top \boldsymbol{x} \geq \beta^q_{\boldsymbol{\alpha}}(S)$$

  is valid inequality for $\mathcal{X}$.
- Quantile cuts can be useful in solving the MILP formulation of CCP (2)
- More new valid inequalities can be generated from a set of base valid inequalities by mixing techniques

Quantile Closure

- Define the quantile closure

$$S^1 = \bigcap_{\boldsymbol{\alpha} \in \mathbb{R}^n} \{\boldsymbol{x} : \boldsymbol{\alpha}^\top \boldsymbol{x} \geq \beta_q^{\boldsymbol{\alpha}}(S)\}$$

and note $\mathrm{Proj}_x(\mathcal{X}) \subseteq S^1$, hence we can iteratively define

$$S^r = \bigcap_{\boldsymbol{\alpha} \in \mathbb{R}^n} \{\boldsymbol{x} : \boldsymbol{\alpha}^\top \boldsymbol{x} \geq \beta_q^{\boldsymbol{\alpha}}(S^{r-1})\}, \text{ for } r > 1$$

- Assume S is compact, a theoretically interesting result[2] is

$$\lim_{r \to \infty} d_H(S^r, \mathrm{conv}\left(\mathrm{Proj}_x(\mathcal{X})\right)) = 0$$

[2] Weijun Xie and Shabbir Ahmed. "On quantile cuts and their closure for chance constrained optimization problems". In: *Mathematical Programming* 172.1 (2018), pp. 621–646.

CVaR Approximation

Idea: convexify the problem

- Note that for any $y > 0$,

$$\mathbb{P}\left[\tilde{v} > 0\right] = \mathbb{E}_{\mathbb{P}}\left[\mathbb{I}(\tilde{v} > 0)\right] \leq \mathbb{E}_{\mathbb{P}}\left[(\tilde{v}/y + 1)_+\right]$$

So

$$\begin{aligned}
& \mathbb{P}\left[\tilde{v} \leq 0\right] \geq 1 - \epsilon \\
\iff \quad & \mathbb{P}\left[\tilde{v} > 0\right] \leq \epsilon \\
\Longleftarrow \quad & \inf_{y>0} \mathbb{E}_{\mathbb{P}}\left[(\tilde{v}/y + 1)_+\right] \leq \epsilon \\
\iff \quad & \inf_{y>0} \mathbb{E}_{\mathbb{P}}\left[(\tilde{v} + y)_+\right] - y\epsilon \leq 0 \\
\iff \quad & \inf_{y<0} \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[(\tilde{v} - y)_+\right] + y \leq 0 \\
\iff \quad & \inf_{y\in\mathbb{R}} \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[(\tilde{v} - y)_+\right] + y \leq 0 \;\; \text{i.e., } \mathrm{CVaR}_{1-\epsilon}(\tilde{v}) \leq 0
\end{aligned}$$

CVaR Approximation

For joint chance constraints $\mathbb{P}\left[\tilde{\boldsymbol{A}}\boldsymbol{x} \geq \tilde{\boldsymbol{b}}\right] \geq 1 - \epsilon$, we have a safe convex approximation

$$\inf_{y \in \mathbb{R}} \frac{1}{\epsilon} \mathbb{E}_{\mathbb{P}}\left[(\max_{\ell \in [m]} \{\tilde{b}_\ell - \tilde{\boldsymbol{a}}_\ell^\top \boldsymbol{x}\} - y)_+ \right] + y \leq 0$$

CVaR Approximation

For joint chance constraints $\mathbb{P}\left[\tilde{\boldsymbol{A}}\boldsymbol{x} \geq \tilde{\boldsymbol{b}}\right] \geq 1-\epsilon$, we have a safe convex approximation

$$\inf_{y\in\mathbb{R}} \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[(\max_{\ell\in[m]}\{\tilde{b}_\ell - \tilde{\boldsymbol{a}}_\ell^\top \boldsymbol{x}\} - y)_+\right] + y \leq 0$$

So the CCP has an upper bound

$$\min_{\boldsymbol{x},y,\theta_1,\dots,\theta_N} \left\{ \boldsymbol{c}^\top\boldsymbol{x} : \begin{array}{ll} y + \frac{1}{\epsilon}\sum_{i\in[N]} p_i\theta_i \leq 0 & \\ \theta_i \geq 0 & \forall i \in [N] \\ \theta_i \geq b_\ell^i - \boldsymbol{a}_\ell^{i\top}\boldsymbol{x} - y & \forall i\in[N], \ell\in[m] \end{array} \right\}$$

Bad Example

Example 1

Given a parameter $\kappa > 1$, consider

$$v_\kappa^* = \min_{x \geq 0} \left\{ x : \mathbb{P}\left[\tilde{\xi} x \geq 1\right] \geq 1 - \epsilon \right\}$$

where $\mathbb{P}\left[\tilde{\xi} = 1\right] = \epsilon$ and $\mathbb{P}\left[\tilde{\xi} = \kappa\right] = 1 - \epsilon$.

Bad Example

Example 1

Given a parameter $\kappa > 1$, consider

$$v_\kappa^* = \min_{x \geq 0} \left\{ x : \mathbb{P}\left[\tilde{\xi} x \geq 1\right] \geq 1 - \epsilon \right\}$$

where $\mathbb{P}\left[\tilde{\xi} = 1\right] = \epsilon$ and $\mathbb{P}\left[\tilde{\xi} = \kappa\right] = 1 - \epsilon$.

- $v_\kappa^* = 1/\kappa$

Bad Example

Example 1

Given a parameter $\kappa > 1$, consider

$$v_\kappa^* = \min_{x \geq 0} \left\{ x : \mathbb{P}\left[\tilde{\xi} x \geq 1\right] \geq 1 - \epsilon \right\}$$

where $\mathbb{P}\left[\tilde{\xi} = 1\right] = \epsilon$ and $\mathbb{P}\left[\tilde{\xi} = \kappa\right] = 1 - \epsilon$.

- $v_\kappa^* = 1/\kappa$
- $v_\kappa^{CVaR} = 1$

Bad Example

Example 1

Given a parameter $\kappa > 1$, consider

$$v_\kappa^* = \min_{x \geq 0} \left\{ x : \mathbb{P}\left[\tilde{\xi} x \geq 1\right] \geq 1 - \epsilon \right\}$$

where $\mathbb{P}\left[\tilde{\xi} = 1\right] = \epsilon$ and $\mathbb{P}\left[\tilde{\xi} = \kappa\right] = 1 - \epsilon$.

- $v_\kappa^* = 1/\kappa$
- $v_\kappa^{CVaR} = 1$
- So

$$\lim_{\kappa \to \infty} \frac{v_\kappa^{CVaR}}{v_\kappa^*} = \infty$$

Scenario Approximation

Idea: enforce the constraints in sampled scenarios

- Sample $\bar{N}$ i.i.d. samples from $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{b}})$, and compute

$$v^{SA} = \min_{\boldsymbol{x}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{A}^k \boldsymbol{x} \geq \boldsymbol{b}^k \quad \forall k \in [\bar{N}] \right\}$$

[3] Giuseppe Carlo Calafiore and Marco C Campi. "The scenario approach to robust control design". In: *IEEE Transactions on automatic control* 51.5 (2006), pp. 742–753.

Scenario Approximation

Idea: enforce the constraints in sampled scenarios

- Sample $\bar{N}$ i.i.d. samples from $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{b}})$, and compute

$$v^{SA} = \min_{\boldsymbol{x}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{A}^k \boldsymbol{x} \geq \boldsymbol{b}^k \quad \forall k \in [\bar{N}] \right\}$$

- It is justified when $\bar{N}$ is large enough, say[3]

$$\bar{N} \geq \frac{2}{\epsilon} \log\left(\frac{1}{\delta}\right) + \frac{2n}{\epsilon} \log\left(\frac{2}{\epsilon}\right) + 2n,$$

  and suppose $v^{SA} < \infty$, then its solution is also feasible to CCP with probability at least $1 - \delta$.

[3] **Calafiore and Campi**, "The scenario approach to robust control design".

Scenario Approximation

Idea: enforce the constraints in sampled scenarios

- Sample $\bar{N}$ i.i.d. samples from $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{b}})$, and compute

$$v^{SA} = \min_{\boldsymbol{x}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{A}^k \boldsymbol{x} \geq \boldsymbol{b}^k \quad \forall k \in [\bar{N}] \right\}$$

- It is justified when $\bar{N}$ is large enough, say[3]

$$\bar{N} \geq \frac{2}{\epsilon} \log\left(\frac{1}{\delta}\right) + \frac{2n}{\epsilon} \log\left(\frac{2}{\epsilon}\right) + 2n,$$

  and suppose $v^{SA} < \infty$, then its solution is also feasible to CCP with probability at least $1 - \delta$.
- However, v^{SA} can be ∞ with high probability even for simple examples.

[3] **Calafiore and Campi**, "The scenario approach to robust control design".

Bad Example

Example 2

Consider

$$v^* = \min_{x \geq 0} \left\{ x : \mathbb{P}\left[\tilde{\xi} x \geq 1\right] \geq 1 - \epsilon \right\}$$

where $\mathbb{P}\left[\tilde{\xi} = 0\right] = \epsilon$ and $\mathbb{P}\left[\tilde{\xi} = 1\right] = 1 - \epsilon$.

Note that

- $v_* = 1$
- Given confidence parameter δ, let $\bar{N}_\delta = \lceil \frac{2}{\epsilon} \log\left(\frac{1}{\delta}\right) + \frac{2}{\epsilon} \log\left(\frac{2}{\epsilon}\right) + 2 \rceil$, then

  $$\mathbb{P}\left[v_\delta^{SA} = \infty\right] \geq 1 - \delta^2 \epsilon^2$$

  because the SA problem

  $$\min_{x \geq 0} \left\{ x : \xi^i x \geq 1, \forall i \in [\bar{N}_\delta] \right\}$$

  is infeasible whenever there is some $\xi^i = 0$, which happens with probability at least $1 - (1 - \epsilon)^{\bar{N}_\delta} \geq 1 - \delta^2 \epsilon^2$.

Bisection Heuristic

Idea: minimize the expected constraint violation with fixed budget, then bisection to optimize the budget parameter

The bisection heuristic works as follows:

- Fix v and solve

$$\min_{\boldsymbol{x},\boldsymbol{s}} \left\{ \sum_{i\in[N]} p_i s_i : \begin{array}{ll} s_i \boldsymbol{e} \geq \boldsymbol{b}^i - \boldsymbol{A}^i \boldsymbol{x} & \forall i \in [N] \\ s_i \geq 0 & \forall i \in [N] \\ \boldsymbol{c}^\top \boldsymbol{x} \leq v & \end{array} \right\}$$

- Check whether $\sum_{i\in[N]} p_i \mathbb{I}(s_i = 0) \geq 1 - \epsilon$, then decrease v if yes, otherwise increase v.

Bad Example

Example 3

Let $\epsilon = 1 - 1/N$, consider

$$v^* = \min_{\boldsymbol{x} \in \mathbb{R}^2_+} \left\{ \frac{1}{1-\epsilon} x_1 + x_2 : \mathbb{P}\left[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{x} \geq 1\right] \geq 1 - \epsilon \right\}$$

where $\mathbb{P}\left[\tilde{\boldsymbol{\xi}} = (1,0)\right] = \epsilon$ and $\mathbb{P}\left[\tilde{\boldsymbol{\xi}} = (1,1)\right] = 1 - \epsilon$.

Then $v^* = 1$ and optimal $\boldsymbol{x} = (0,1)$.

Bad Example

Example 3

Let $\epsilon = 1 - 1/N$, consider

$$v^* = \min_{\boldsymbol{x} \in \mathbb{R}^2_+} \left\{ \frac{1}{1-\epsilon} x_1 + x_2 : \mathbb{P}\left[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{x} \geq 1\right] \geq 1 - \epsilon \right\}$$

where $\mathbb{P}\left[\tilde{\boldsymbol{\xi}} = (1,0)\right] = \epsilon$ and $\mathbb{P}\left[\tilde{\boldsymbol{\xi}} = (1,1)\right] = 1 - \epsilon$.

Then $v^* = 1$ and optimal $\boldsymbol{x} = (0,1)$.

However, start from $u > \frac{1}{1-\epsilon}$ an $L = 0$, for any $v \in [1, \frac{1}{1-\epsilon})$, the bisection subproblem

$$\min_{\boldsymbol{x} \in \mathbb{R}^2_+, \boldsymbol{s} \in \mathbb{R}^N_+} \left\{ \frac{1}{N} \sum_{i \in [N]} s_i : \begin{array}{ll} s_i \geq 1 - x_1 & \forall i \in [\epsilon N] \\ s_i \geq 1 - x_1 - x_2 & \forall i \in [N] \backslash [\epsilon N] \\ \frac{1}{1-\epsilon} x_1 + x_2 \leq v & \end{array} \right\}$$

has optimal solution $\boldsymbol{x} = ((1-\epsilon)v, 0)$ and $s_i = 1 - (1-\epsilon)v > 0$, so that $v^H \geq \frac{1}{1-\epsilon} = N$.

Bicriteria Approximation

Definition 1

Given a violation ratio $\sigma \geq 1$ and an approximation ratio $\gamma \geq 1$, a (σ, γ)-bicriteria approximation algorithm for CCP returns a solution $(\boldsymbol{x}, \boldsymbol{z}) \in \mathbb{R}^n \times \{0,1\}^N$ such that

$$\begin{aligned} &\boldsymbol{A}^i \boldsymbol{x} \geq \boldsymbol{b}^i z_i - (1 - z_i)\boldsymbol{M}_i, \quad \forall i \in [N] \\ &\textstyle\sum_{i \in [N]} p_i z_i \geq 1 - \sigma\epsilon, \quad \boldsymbol{c}^\top \boldsymbol{x} \leq \gamma v^* \end{aligned}$$

An inapproximability result is

Theorem 1

Suppose we have a polynomial time algorithm that returns a (σ, γ) approximate solution to the CCP with $p_i = 1/N$. Then, unless P=N P, the following holds:

1. *If $\gamma = 1$, then we must have $\sigma = 1/\epsilon - f(N)(1-\epsilon)/\epsilon$ for some function f such that $f(N) \to 0$ as $N \to \infty$*
2. *If $\sigma = 1$, then we must have $\gamma = g(N)$ for some function g such that $g(N) \to \infty$ as $N \to \infty$.*

Relax-and-Scale Algorithm

For simplicity, consider a special CCP where the big-M reduce to zero,

$$v^* = \min_{\boldsymbol{x}\in\mathbb{R}^n_+, \boldsymbol{z}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \boldsymbol{A}^i \boldsymbol{x} \geq z_i \boldsymbol{b}^i \\ \sum_{i\in[N]} p_i z_i \geq 1-\epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\} \qquad \text{(Covering CCP)}$$

where $\boldsymbol{A}^i \in \mathbb{R}^{m\times n}_+, \boldsymbol{b}^i \in \mathbb{R}^m_+, \boldsymbol{c} \in \mathbb{R}^n_+$.

Relax-and-Scale Algorithm

For simplicity, consider a special CCP where the big-M reduce to zero,

$$v^* = \min_{\boldsymbol{x}\in\mathbb{R}^n_+,\boldsymbol{z}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \boldsymbol{A}^i \boldsymbol{x} \geq z_i \boldsymbol{b}^i \\ \sum_{i\in[N]} p_i z_i \geq 1-\epsilon \\ \boldsymbol{z} \in \{0,1\}^N \end{array} \right\} \qquad \text{(Covering CCP)}$$

where $\boldsymbol{A}^i \in \mathbb{R}^{m\times n}_+, \boldsymbol{b}^i \in \mathbb{R}^m_+, \boldsymbol{c} \in \mathbb{R}^n_+$.

Algorithm 2: Relax-and-Scale (σ, γ)

Input: paramter $\sigma \geq 1$, $\gamma \geq 1$
Initialization Set $L = 1$, $U = \gamma$, and stopping tolerance $\delta > 0$;
Solve LP relaxation of CCP and denote its optimal solution $(\boldsymbol{x}, \boldsymbol{z})$
while $U - L > \delta$ **do**
  Set $\tau = (L+U)/2$;
  Set $\boldsymbol{x} \leftarrow \tau \boldsymbol{x}$, $\boldsymbol{z} \leftarrow \min\{\lfloor \tau \boldsymbol{z}, \boldsymbol{e} \rfloor$;
  Check whether $\sum_{i\in[N]} p_i z_i \geq 1 - \sigma\epsilon$, then set $U \leftarrow \tau$ if yes,
   otherwise set $L \leftarrow \tau$;
end
Output: $\bar{x} \leftarrow U\boldsymbol{x}$

Bicriteria Analysis

Theorem 2

Assume $p_i = 1/N$*, suppose we choose*

$$\sigma \in [1, 1/\epsilon), \quad \gamma = \frac{1 + \lfloor \sigma \epsilon N \rfloor}{1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N},$$

then the Relax-and-Scale algorithm returns a (σ, γ)*-bicriteria approximate solution. Furthermore,* $\gamma \leq \frac{\sigma}{\sigma - 1}$.

Bicriteria Analysis

Theorem 2

Assume $p_i = 1/N$, *suppose we choose*

$$\sigma \in [1, 1/\epsilon), \quad \gamma = \frac{1 + \lfloor \sigma \epsilon N \rfloor}{1 + \lfloor \sigma \epsilon N \rfloor - \epsilon N},$$

then the Relax-and-Scale algorithm returns a (σ, γ)*-bicriteria approximate solution. Furthermore,* $\gamma \leq \frac{\sigma}{\sigma - 1}$.

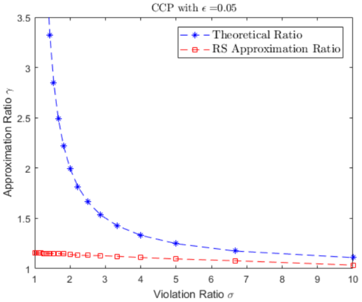
The bound could be conservative when $\sigma \to 1$ in practice.
Solve a CCP

$$v^* = \min_{\boldsymbol{x} \in \mathbb{R}^{50}_+, \boldsymbol{z} \in \{0,1\}^{100}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{\xi}^{i\top} \boldsymbol{x} \geq z_i, \forall i \in [100], \sum_{i \in [100]} z_i / 100 \geq (1 - \epsilon) \right\}$$

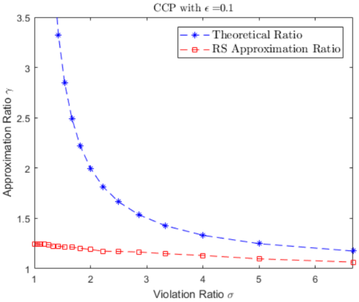
where ξ^i are i.i.d. uniformly sampled between 0.8 to 1.2 and $\epsilon \in \{0.05, 0.10\}$.

Numerical Illustration

Theoretical ratio: $\sigma/(\sigma-1)$



(a) $\epsilon = 0.05$



(b) $\epsilon = 0.10$

Suggest choosing $\sigma \in (1, 2]$.

Summary

- CCP is a challenging optimization model which motivates many useful approximation techniques
- CCP is ill-posed in some sense, be aware of its distributionally robust counterparts and other alternatives such as risk or robust constraints[4]
- Combine different approximation techniques can be useful, e.g., bisection + CVaR approximation[5]
- Solution techniques from CCP can be useful for solving other problems such as cardinality constrained problems[6]

[4] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. ***Robust optimization***. Princeton university press, 2009.

[5] Jiang Nan and Xie Weijun. "ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs.". In: *Operations Research* (2021).

[6] Weijun Xie and Xinwei Deng. "Scalable algorithms for the sparse ridge regression". In: *SIAM Journal on Optimization* 30.4 (2020), pp. 3359–3386.