CVaR, Uncertainty Set, and the Joint Safeguarding Constraint[1]

J. Sun
National University of Singapore

April, 2008

__Abstract__ The condition-value-at-risk (CVaR) studied by Rockafellar and Uryasev could be used to approximate a single safeguarding constraint in stochastic optimization. There is a natural connection between a computationally tractable approximation to a CVaR constraint and an optimization problem with respect to an uncertainty set. This idea is extended to handling the joint safeguarding constraint problem, resulting practical improvement upon the current approach.

This talk is originated from a joint work with Chen, Sim, and Teo at National University of Singapore.

[1]The main part of the talk is from a joint paper with W. Chen, M. Sim, and C. Teo at National University of Singapore, School of Business.

1 Conditional Value at Risk (CVaR) and Stochastic Constraints

Consider a "stochastic" constraint

$$f(x, \tilde{z}) \leq 0 \tag{1}$$

A way of safeguarding x from being infeasible is to assign a probability such as

$$\boldsymbol{P}(f(x, \tilde{z}) \leq 0) \geq 1 - \varepsilon. \tag{2}$$

Great model, terrible to compute!

- Involving multi-dimensional integrals
- Producing non-convexity

Starting from the simplest (but widely applicable) case

$$f(x, \tilde{z}) \leq 0 \quad \Rightarrow \quad \boldsymbol{a}(\tilde{z})'x \leq b(\tilde{z}),$$

where

$$\begin{aligned} \boldsymbol{a}(\tilde{\boldsymbol{z}}) &= a^0 + \sum_{k=1}^{N} \boldsymbol{a}^k \tilde{z}_k \\ b(\tilde{\boldsymbol{z}}) &= b_0 + \sum_{k=1}^{N} b_k \tilde{z}_k. \end{aligned}$$

Then

$$\boldsymbol{a}(\tilde{z})'x \leq b(\tilde{z}) \quad \Longleftrightarrow \quad y(\tilde{z}) \leq 0,$$

where

$$y(\tilde{z}) = y_0 + \sum_{i=1}^{N} y_k \tilde{z}_k \quad \text{and} \quad y_k = (\boldsymbol{a}^k)'x - b_k, \ \forall k = 0, ..., N$$

Thus, here and below, we concentrate on the probabilistic (chance) constraint

$$\boldsymbol{P}(y(\tilde{z}) \leq 0) \geq 1 - \varepsilon, \tag{3}$$

which is equivalent to

$$\mathrm{VaR}_{1-\varepsilon}(y(\tilde{z})) \leq 0.$$

A step towards tractability is by convexifying the constraint (3) using the CVaR as proposed by Rockafellar and Uryasev.

$$\mathrm{CVaR}_{1-\varepsilon}(y(\tilde{z})) \leq 0 \quad \Rightarrow \quad \mathrm{VaR}_{1-\varepsilon}(y(\tilde{z})) \leq 0,$$

where

$$\mathrm{CVaR}_{1-\varepsilon}(y(\tilde{z})) \;\triangleq\; \min_{\beta} \left\{ \beta + \frac{1}{\epsilon} \boldsymbol{E}\left[(y(\tilde{z}) - \beta\right]^{+} \right\}.$$

- Convex in $y = (y_0, y_1, ..., y_N)' \triangleq (y_0, \boldsymbol{y})'$;
- Still need to compute a multi-dimensional integral except for discrete distribution.

Idea: Find an upper bound for $\boldsymbol{E}\left[(y(\tilde{z}) - \beta\right]^{+}$, which is easier to compute and substitute CVaR by

$$\eta_{1-\epsilon}(y_0, \boldsymbol{y}) \triangleq \min_{\beta} \left\{ \beta + \frac{1}{\epsilon} \pi(y_0 - \beta, \boldsymbol{y}) \right\}. \tag{4}$$

Several good (tight) bounds for $\boldsymbol{E}\left[(y(\tilde{z})-\beta\right]^+$:

1. $\pi^1(y_0-\beta,\boldsymbol{y}) \triangleq \left(y_0-\beta+\max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{z}'\boldsymbol{y}\right)^+$, where $\mathcal{W}$ is a box containing the support of the distribution.
2. $\pi^2(y_0-\beta,\boldsymbol{y}) \triangleq y_0+\left(-y_0+\beta+\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z}\right)^+$.
3. $\pi^3(y_0-\beta,\boldsymbol{y}) \triangleq \frac{1}{2}(y_0-\beta)+\frac{1}{2}\sqrt{(y_0-\beta)^2+\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}$.
4. $\pi^4(y_0-\beta,\boldsymbol{y}) \triangleq \inf_{\mu>0}\{\frac{\mu}{e}\exp(\frac{y_0-\beta}{\mu}+\frac{\|\boldsymbol{u}\|_2^2}{2\mu^2})\}$, where $u_j=\max\{p_jy_j,-q_jy_j\}$, $j=1,\ldots,N$.
5. $\pi^5(y_0-\beta,\boldsymbol{y}) \triangleq y_0+\inf_{\mu>0}\{\frac{\mu}{e}\exp(-\frac{y_0-\beta}{\mu}+\frac{\|\boldsymbol{v}\|_2^2}{2\mu^2})\}$,where $v_j=\max\{-p_jy_j,q_jy_j\}$, $j=1,\ldots,N$.

By replacing $\boldsymbol{E}\left[(y(\tilde{z})-\beta\right]^+$ with $\pi^i(y_0-\beta,\boldsymbol{y})$, we safeguard the CVaR constraint by the constraint

$$\eta_{1-\epsilon}(y_0,\boldsymbol{y}) \triangleq \min_{\beta}\left\{\beta+\frac{1}{\epsilon}\pi(y_0-\beta,\boldsymbol{y})\right\}\le 0, \tag{5}$$

where $\pi(\cdot,\cdot)$ is a positively homogeneous, closed, proper convex function. We assume in the following that the uncertainties $\{\tilde{z}_j\}_{j=1:N}$ are zero mean random variables.

Theorem 1.1 *Suppose that $\pi(y_0,\boldsymbol{y})$ is convex, closed, and positively homogeneous, and is an upper bound to $\boldsymbol{E}\left((y_0+\boldsymbol{y}'z)^+\right)$ with $\pi(y_0,\boldsymbol{0})=y_0^+$. Then it holds that for all $(y_0,\boldsymbol{y})$ such that $\pi(y_0,\boldsymbol{y})<\infty$, we have*

$$\eta_{1-\epsilon}(y_0,\boldsymbol{y}) = y_0+\max_{\boldsymbol{z}\in\mathcal{U}(\epsilon)}\boldsymbol{y}'\boldsymbol{z}$$

for some convex uncertainty set $\mathcal{U}(\epsilon)$.

This theorem says that the η function is second-order-cone representable. Therefore, we

- avoid multi-dimensional integrals and
- preserve the convexity.

There is a 1-1 correspondence

$$\text{The upper bound } \eta_{1-\varepsilon} \text{ to CVaR} \quad \Longleftrightarrow \quad \text{The uncertainty set } \mathcal{U}(\epsilon).$$

The uncertainty sets corresponding to $\eta^{i}_{1-\varepsilon}(y_0, \boldsymbol{y})$, $i = 1, \ldots, 5$ are as follows.

$$\begin{aligned}
\mathcal{U}_1(\epsilon) &\triangleq \quad \mathcal{W}, \\
\mathcal{U}_2(\epsilon) &\triangleq \quad \{\boldsymbol{z} \mid \boldsymbol{z} = (1 - 1/\epsilon)\boldsymbol{\zeta}, \text{for some } \boldsymbol{\zeta} \in \mathcal{W}\}, \\
\mathcal{U}_3(\epsilon) &\triangleq \quad \left\{\boldsymbol{z} \mid \|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{z}\|_2 \leq \sqrt{\frac{1-\epsilon}{\epsilon}}\right\}, \\
\mathcal{U}_4(\epsilon) &\triangleq \quad \left\{\boldsymbol{z} \mid \exists \boldsymbol{s}, \boldsymbol{t} \in \Re^N, \boldsymbol{z} = \boldsymbol{s} - \boldsymbol{t}, \|\boldsymbol{P}^{-1}\boldsymbol{s} + \boldsymbol{Q}^{-1}\boldsymbol{t}\|_2 \leq \sqrt{-2\ln\epsilon}\right\}, \\
\mathcal{U}_5(\epsilon) &\triangleq \quad \left\{\boldsymbol{z} \mid \exists \boldsymbol{s}, \boldsymbol{t} \in \Re^N, \boldsymbol{z} = \boldsymbol{s} - \boldsymbol{t}, \|\boldsymbol{Q}^{-1}\boldsymbol{s} + \boldsymbol{P}^{-1}\boldsymbol{t}\|_2 \leq \frac{1-\epsilon}{\epsilon}\sqrt{-2\ln(1-\epsilon)}\right\}.
\end{aligned}$$

where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are diagonal matrices involving “forward” and “backward” deviations.

2 The Joint Chance Constrained Problem

We consider a linear joint chance constraint as follows,

$$\mathrm{P}(y_j(\tilde{\boldsymbol{z}}) \leq 0, j \in \mathcal{M}) \geq 1 - \epsilon, \tag{6}$$

where $\mathcal{M} = \{1, \ldots, m\}$, $y_j(\tilde{\boldsymbol{z}})$ are affinely dependent on $\tilde{\boldsymbol{z}}$,

$$y_j(\tilde{\boldsymbol{z}}) = y_j^0 + \sum_{k=1}^{N} y_j^k \tilde{z}_k \qquad j \in \mathcal{M}.$$

$(y_1^0, \ldots, y_1^N, \ldots, y_m^0, \ldots, y_m^N)$ being the decision variables. For notational convenience, we represent

$$\boldsymbol{y}_j = (y_j^1, \ldots, y_j^N),$$

so that $y_i(\tilde{\boldsymbol{z}}) = y_i^0 + \boldsymbol{y}_i' \tilde{\boldsymbol{z}}$ and denote

$$\boldsymbol{Y} = (y_1^0, \ldots, y_1^N, \ldots, y_m^0, \ldots, y_m^N),$$

as the collection of decision variables in the joint chance constraint.

We could treat the joint chance constraint as combination of single ones since

$$\mathrm{P}(y_i(\tilde{\boldsymbol{z}}) \leq 0) \geq 1 - \epsilon_i,\ i \in \mathcal{M}, \sum_{i \in \mathcal{M}} \epsilon_i \leq \epsilon \Rightarrow \mathrm{P}(y_j(\tilde{\boldsymbol{z}}) \leq 0, j \in \mathcal{M}) \geq 1 - \epsilon,$$

but it is often too conservative, and there is no guidance of how to select ε_i.

We propose a new tractable way for approximating the joint chance constraint. Given a vector of positive weights, $\sum \boldsymbol{\alpha}_i = 1$, $\boldsymbol{\alpha} > \boldsymbol{0}$, an upper bound $\pi(y_0, \boldsymbol{y})$ for $\boldsymbol{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right)$, we define the following function,

$$\gamma_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}) \triangleq \min_{w_0, \tilde{\boldsymbol{w}}} \left(\eta_{1-\epsilon}(w_0, \boldsymbol{w}) + \frac{1}{\epsilon} \left\{ \sum_{i \in \mathcal{M}} \pi(\alpha_i y_i^0 - w_0, \alpha_i \boldsymbol{y}_i - \boldsymbol{w}) \right\} \right).$$

The next result shows we can use the above function to approximate a joint chance constraint.

Theorem 2.1 (a)

$$\mathrm{CVaR}_{1-\epsilon}\left(\max_{i\in\mathcal{M}}\{\alpha_i y_i(\tilde{\boldsymbol{z}})\}\right) \leq \gamma_{1-\epsilon}(\boldsymbol{Y},\boldsymbol{\alpha}).$$

Consequently, the joint chance constraint (6) is satisfied if

$$\gamma_{1-\epsilon}(\boldsymbol{Y},\boldsymbol{\alpha}) \leq 0. \tag{7}$$

(b) *For fixed $\boldsymbol{\alpha}$, the epigraph of the function $\gamma_{1-\epsilon}(\boldsymbol{Y},\boldsymbol{\alpha})$ with respect to $\boldsymbol{Y}$ is second-order cone representable and positive homogeneous. Similarly, for a fixed $\boldsymbol{Y}$, the epigraph of the function $\gamma_{1-\epsilon}(\boldsymbol{Y},\boldsymbol{\alpha})$ with respect to $\boldsymbol{\alpha}$ is second-order cone representable and positive homogeneous.*

The proof of this theorem is based on a classical inequality

$$\boldsymbol{E}\left(\max_{i=1,\ldots,n} X_i - \beta\right)^+ \leq \boldsymbol{E}\left(Y-\beta\right)^+ + \sum_{i=1}^{n}\boldsymbol{E}\left(X_i - Y\right)^+, \text{ for any r.v. } Y.$$

Unfortunately, γ is not jointly convex in $(\boldsymbol{Y},\boldsymbol{\alpha})$. However, we can alternatively optimize $\boldsymbol{Y}$ and $\boldsymbol{\alpha}$ with the hope to find certain pair $(\boldsymbol{Y},\boldsymbol{\alpha})$ such that (7) is valid. Indeed, under mild assumptions, the algorithm converges to such a solution. Details are omitted.

3 Computational Studies

We consider a resource allocation problem on a network in which the demands are uncertain. The network we consider is an directed graph with node set $\mathcal{V}$, $|\mathcal{V}| = n$ and arc set $\mathcal{E}$, $|\mathcal{E}| = r$. At each node, i, $i \in \mathcal{V}$, we decide on the quantity of resource x_i to allocate, which will incur a cost of c_i per unit resource. When the demands $\tilde{d}_i$, $i \in \mathcal{V}$ are realized, resources at the nodes or from neighboring nodes are used to meet the demands. The goal is to minimize the total allocation cost subjected to a service level constraint of meeting all demands with probability at least $1 - \epsilon$. We assume that the resource at each node i can only be transshipped across to its outgoing neighboring nodes defined as

$$\mathcal{N}^{-}(i) \triangleq \{j : (i,j) \in \mathcal{E}\},$$

and received from its incoming neighboring nodes defined as

$$\mathcal{N}^{+}(i) \triangleq \{j : (j,i) \in \mathcal{E}\}.$$

Transshipment of resources received from other nodes is prohibited.

In our model, we ignore operating costs such as the transshipment costs. One of such applications is with regards to allocation of equipment such as ambulances or time critical medical supplies for emergency response to local or neighboring demands. The costs associated with their procurement is more significant than the operating cost of transshipment, which may occur rather infrequently. We list the notations of the model as follows

- c_i : Unit cost of having one resource at node i, $i \in \mathcal{V}$;
- $d_i(\tilde{\boldsymbol{z}})$: Demand at node i, $i \in \mathcal{V}$ as a function of the primitive uncertainties $\tilde{\boldsymbol{z}}$;
- x_i: Quantity of resource at node i, $i \in \mathcal{V}$;

- $w_{ij}(\tilde{\boldsymbol{z}})$: Transshipment quantity from node i to node j, $(i,j) \in \mathcal{E}$ in respond to realization of $\tilde{\boldsymbol{z}}$.

The problem can be formulated as a joint chance constrained problem as follows,

$$
\begin{array}{ll}
\min & \boldsymbol{c}'\boldsymbol{x} \\
\text{s.t.} & \boldsymbol{P}\left(\begin{array}{ll}
x_i + \sum\limits_{j\in\mathcal{N}^+(i)} w_{ji}(\tilde{\boldsymbol{z}}) - \sum\limits_{j\in\mathcal{N}^-(i)} w_{ij}(\tilde{\boldsymbol{z}}) \geq d_i(\tilde{\boldsymbol{z}}) & i=1,\ldots,n \\
x_i \geq \sum\limits_{j\in\mathcal{N}^-(i)} w_{ij}(\tilde{\boldsymbol{z}}) & i=1,\ldots,n \\
\boldsymbol{w}(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0} &
\end{array}\right) \geq 1-\epsilon \\
& \boldsymbol{x} \geq \boldsymbol{0}.
\end{array} \tag{8}
$$

We assume that the demand at each node are independently distributed and represented as

$$
d_j(\tilde{\boldsymbol{z}}) = d_j^0 + \tilde{z}_j,
$$

where $\tilde{z}_j$ are independent zero mean random variables with unknown distribution.

By introducing new variables, we can transform the model (8) to the "standard form" model as follows

$$
\begin{array}{lll}
\min & \boldsymbol{c}'\boldsymbol{x} & \\
\text{s.t.} & x_i + \sum\limits_{j\in\mathcal{N}^+(i)} w_{ji}(\tilde{\boldsymbol{z}}) - \sum\limits_{\in\mathcal{N}^-(i)} w_{ij}(\tilde{\boldsymbol{z}}) + \boldsymbol{r}(\tilde{\boldsymbol{z}}) = d_i(\tilde{\boldsymbol{z}}) & i=1,\ldots,n \\
& x_i + s_i(\tilde{\boldsymbol{z}}) = \sum\limits_{j\in\mathcal{N}^-(i)} w_{ij}(\tilde{\boldsymbol{z}}) & i=1,\ldots,n \\
& \boldsymbol{w}(\tilde{\boldsymbol{z}}) + \boldsymbol{t}(\tilde{\boldsymbol{z}}) = \boldsymbol{0} & \\
& \boldsymbol{y}(\tilde{\boldsymbol{z}}) = \begin{pmatrix} \boldsymbol{r}(\tilde{\boldsymbol{z}}) \\ \boldsymbol{s}(\tilde{\boldsymbol{z}}) \\ \boldsymbol{t}(\tilde{\boldsymbol{z}}) \end{pmatrix} & \\
& \boldsymbol{P}(\boldsymbol{y}(\tilde{\boldsymbol{z}}) \leq \boldsymbol{0}) \geq 1-\epsilon & \\
& \boldsymbol{x} \geq \boldsymbol{0}. &
\end{array} \tag{9}
$$

Note that the dimension of $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ is $m = 2n + r$.

In our test problem, we generate 15 nodes randomly positioned on a square grid and restrict to the r shortest arcs on the grid in terms of Euclidean distances. We assume $c_i = 1$. For the demand uncertainty, we assume that $d_j^0 = 10$ and the demand at each node, $d_j(\tilde{\boldsymbol{z}})$ takes value from zero to 100.

We first solve the problem using the classical approach by decomposing the joint chance constrained problem into m constraints of the form

$$\eta_{1-\epsilon_i}(y_i^0, \boldsymbol{y}_i) \leq 0, \ i \in \mathcal{M}. \tag{10}$$

with $\epsilon_i = \epsilon/(2n + r)$. We denote the optimal solution as $\boldsymbol{x}^B$ and its objective as Z^B. Subsequently, we use the proposed algorithm to improve upon the solution. We report results at the end of twenty iterations. Here, we denote the optimal solution as $\boldsymbol{x}^N$ and its objective as Z^N.

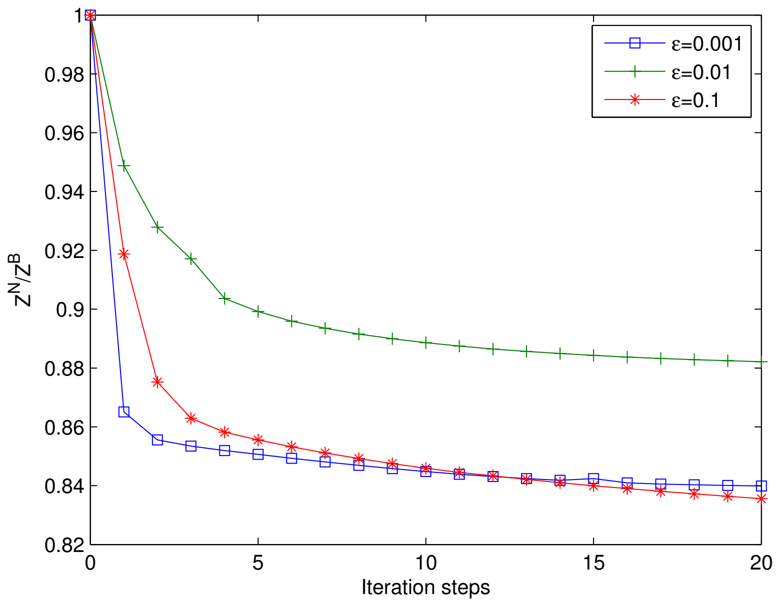

Figure 1: A sample convergence plot