

Stochastic Optimization with Ambiguity in Distribution

J. Sun

Curtin University
and
National University of Singapore

SPCOM 2015, University of South Australia, Adelaide

1 How to Deal with a Single Stochastic Constraint

Consider a single constraint under uncertainty

$$v(x, \tilde{z}) \leq w, \quad (x, w) \in \mathbb{R}^n \times \mathbb{R}$$

Robust Optimization

$$\sup_{z \in C} v(x, z) \leq w \quad (1)$$

Stochastic Optimization

$$\mathbb{E}_{\mathbb{P}_0} v(x, \tilde{z}) \leq w, \quad \mathbb{P}_0 \text{ is the joint distribution of } \tilde{z} \quad (2)$$

Special case of (2) - Chance constraint

$$\mathbb{P}(a(\tilde{z})^T x \geq b(\tilde{z})) \geq 1 - \epsilon \iff \mathbb{E}(\mathbf{1}_{a(\tilde{z})^T x \leq b(\tilde{z})}) \leq \epsilon$$

Sto. Opt . with Ambiguity in Distribution

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} v(x, \tilde{z}) \leq w, \quad \mathcal{P} \text{ is the ambiguity set of distributions.} \quad (3)$$

(1) and (2) \subset (3).

Notations.

$\mathcal{P}_0(\mathbb{R}^m)$ - The set of distributions on \mathbb{R}^m

$|\tilde{z}|, |\tilde{u}|$ - cardinality of \tilde{z} and \tilde{u} , resp.

$\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{|\tilde{z}|} \times \mathbb{R}^{|\tilde{u}|})$ is a joint probability distribution of (\tilde{z}, \tilde{u}) .

$\prod_{\tilde{z}} \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{|\tilde{z}|})$ - the marginal distribution of \tilde{z} under \mathbb{P}

Assumption. $v(x, \tilde{z})$ is convex in x and measurable in $\tilde{z} \forall x$

Let \mathcal{K} be a proper cone, we use $x \succeq_{\mathcal{K}} 0$ to represent $x \in \mathcal{K}$.

1. $\mathcal{K} = \mathbb{R}_+^n \iff x_1, \dots, x_n \geq 0.$

2. $\mathcal{K} = \mathbb{L}^{n+1} \iff x_0 \geq (x_1^2 + \dots + x_n^2)^{1/2}.$

3. $\mathcal{K} = \mathbb{S}_+^n \iff X$ is a positive semidefinite matrix.

More Notations.

$$\mathcal{K}^* := \{y : \langle y, x \rangle \geq 0 \forall x \in \mathcal{K}\}.$$

$$B \preceq_{\mathcal{K}} A \text{ or } A \succeq_{\mathcal{K}} B \iff A - B \in \mathcal{K}.$$

We assume that the *ambiguity set* \mathcal{P} in (3) is representable in the “standard form”

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^P \times \mathbb{R}^Q) : \begin{array}{l} \mathbb{E}_{\mathbb{P}} [A\tilde{z} + B\tilde{u}] = b, \\ \mathbb{P} [(\tilde{z}, \tilde{u}) \in \mathcal{C}_i] \in [\underline{p}_i, \bar{p}_i] \quad \forall i \in \mathcal{I} \end{array} \right\}, \quad (4)$$

where \mathbb{P} represents a joint probability distribution of the random vector $\tilde{z} \in \mathbb{R}^P$ appearing in the constraint function v in (3) and some auxiliary random vector $\tilde{u} \in \mathbb{R}^Q$. We assume that $A \in \mathbb{R}^{K \times P}$, $B \in \mathbb{R}^{K \times Q}$, $b \in \mathbb{R}^K$ and $\mathcal{I} = \{1, \dots, I\}$, while the confidence sets \mathcal{C}_i are defined as

$$\mathcal{C}_i = \{(z, u) \in \mathbb{R}^P \times \mathbb{R}^Q : C_i z + D_i u \preceq_{\mathcal{K}_i} c_i\} \quad (5)$$

with $C_i \in \mathbb{R}^{L_i \times P}$, $D_i \in \mathbb{R}^{L_i \times Q}$, $c_i \in \mathbb{R}^{L_i}$ and \mathcal{K}_i being proper cones.

More requirements on \mathcal{P}

We require that the ambiguity set \mathcal{P} satisfies the following two regularity conditions (**support-interior assumption**).

- (C1) The confidence set \mathcal{C}_I is bounded and has probability one, that is, $\underline{p}_I = \bar{p}_I = 1$.
- (C2) There is a distribution $\mathbb{P} \in \mathcal{P}$ such that $\mathbb{P}[(\tilde{z}, \tilde{u}) \in \mathcal{C}_i] \in (\underline{p}_i, \bar{p}_i)$ whenever $\underline{p}_i < \bar{p}_i$, $i \in \mathcal{I}$.

We require that the constraint function v in (3) satisfies the following condition (**piecewise bi-linear assumption**).

(C3) The constraint function $v(x, z)$ can be written as

$$v(x, z) = \max_{l \in \mathcal{L}} v_l(x, z) \quad (6)$$

where $\mathcal{L} = \{1, \dots, L\}$ and the auxiliary functions $v_l : \mathbb{R}^N \times \mathbb{R}^P \rightarrow \mathbb{R}$ are of the form

$$v_l(x, z) = s_l(z)^\top x + t_l(z) \quad (7)$$

with $s_l(z) = S_l z + s_l$, $S_l \in \mathbb{R}^{N \times P}$ and $s_l \in \mathbb{R}^N$, $t_l(z) = t_l^\top z + t_l$, $t_l \in \mathbb{R}^P$ and $t_l \in \mathbb{R}$.

The tractability of optimization problems with constraints of the type (3) critically depends on the following **nesting condition** for the confidence sets in the definition of \mathcal{P} :

(N) For all $i, i' \in \mathcal{I}$, $i \neq i'$, we have either $\mathcal{C}_i \subseteq \mathcal{C}_{i'}$, $\mathcal{C}_{i'} \subseteq \mathcal{C}_i$ or $\mathcal{C}_i \cap \mathcal{C}_{i'} = \emptyset$.

we denote by $\mathcal{A}(i) = \{i\} \cup \{i' \in \mathcal{I} : \mathcal{C}_i \subseteq \mathcal{C}_{i'}\}$ and $\mathcal{D}(i) = \{i' \in \mathcal{I} : \mathcal{C}_{i'} \subseteq \mathcal{C}_i\}$ the index sets of all supersets (antecedents) and all strict subsets (descendants) of \mathcal{C}_i , respectively.

Theorem 1.1 (Wiesemann, Kuhn, Sim, 2014) *Assume that the conditions (C1)-(C3) and (N) hold. Then, the constraint (3) is satisfied for the ambiguity set (4) if and only if there is $\beta \in \mathbb{R}^K$, $\kappa, \lambda \in \mathbb{R}_+^I$ and $\phi_{il} \in \mathcal{K}_i^*$, $i \in \mathcal{I}$ and $l \in \mathcal{L}$, that satisfy the constraint system*

$$\begin{aligned}
 & \mathbf{b}^\top \boldsymbol{\beta} + \sum_{i \in \mathcal{I}} \left[\bar{p}_i \kappa_i - \underline{p}_i \lambda_i \right] \leq w, \\
 & \left. \begin{aligned}
 & \mathbf{c}_i^\top \phi_{il} + \mathbf{s}_l^\top \mathbf{x} + t_l \leq \sum_{i' \in \mathcal{A}(i)} [\kappa_{i'} - \lambda_{i'}] \\
 & \mathbf{C}_i^\top \phi_{il} + \mathbf{A}^\top \boldsymbol{\beta} = \mathbf{S}_l^\top \mathbf{x} + t_l \\
 & \mathbf{D}_i^\top \phi_{il} + \mathbf{B}^\top \boldsymbol{\beta} = 0
 \end{aligned} \right\} \forall i \in \mathcal{I}, \forall l \in \mathcal{L}.
 \end{aligned} \tag{8}$$

Theorem 1.2 (Lifting Theorem of WKS) *Let $f \in \mathbb{R}^M$ and $g : \mathbb{R}^P \rightarrow \mathbb{R}^M$ be a function with a conic representable \mathcal{K} -epigraph, and consider the ambiguity set*

$$\mathcal{P}' = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^P) : \begin{array}{l} \mathbb{E}_{\mathbb{P}} [g(\tilde{z})] \preceq_{\mathcal{K}} f, \\ \mathbb{P} [\tilde{z} \in \mathcal{C}_i] \in [\underline{p}_i, \bar{p}_i] \quad \forall i \in \mathcal{I} \end{array} \right\}, \quad (9)$$

as well as the lifted ambiguity set

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^P \times \mathbb{R}^M) : \begin{array}{l} \mathbb{E}_{\mathbb{P}} [\tilde{u}] = f, \\ \mathbb{P} [g(\tilde{z}) \preceq_{\mathcal{K}} \tilde{u}] = 1, \\ \mathbb{P} [\tilde{z} \in \mathcal{C}_i] \in [\underline{p}_i, \bar{p}_i] \quad \forall i \in \mathcal{I} \end{array} \right\}, \quad (10)$$

which involves the auxiliary random vector $\tilde{u} \in \mathbb{R}^M$. We then have that (i) $\mathcal{P}' = \Pi_{\tilde{z}} \mathcal{P}$ and (ii) \mathcal{P} can be reformulated as an instance of the standardized ambiguity set (4).

Example 1.1 (Mean) Assume that $G \mathbb{E}_{\mathbb{Q}^0} [\tilde{z}] \preceq_{\mathcal{K}} f$ for a proper cone \mathcal{K} and $G \in \mathbb{R}^{M \times P}$, $f \in \mathbb{R}^M$, and consider the following instance of the ambiguity set (4), which involves the auxiliary random vector $\tilde{u} \in \mathbb{R}^M$.

$$\mathcal{P} = \{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^P \times \mathbb{R}^M) : \mathbb{E}_{\mathbb{P}} [\tilde{u}] = f, \mathbb{P} [G\tilde{z} \preceq_{\mathcal{K}} \tilde{u}] = 1 \} \quad (11)$$

We then have $\mathbb{Q}^0 \in \Pi_{\tilde{z}} \mathcal{P} = \{ \mathbb{Q} \in \mathcal{P}_0(\mathbb{R}^P) : G \mathbb{E}_{\mathbb{Q}} [\tilde{z}] \preceq_{\mathcal{K}} f \}$.

Example 1.2 (Variance) Assume that $\mu = \mathbb{E}_{\mathbb{Q}^0} [\tilde{z}]$ and $\mathbb{E}_{\mathbb{Q}^0} \left[(\tilde{z} - \mu) (\tilde{z} - \mu)^\top \right] \preceq \Sigma$ for a given $\Sigma \in \mathbb{S}_+^P$. Consider the following instance of (4), which involves the auxiliary random matrix $\tilde{U} \in \mathbb{R}^{P \times P}$.

$$\mathcal{P} = \left\{ \mathbb{P} \in (\mathbb{R}^P \times \mathbb{R}^{P \times P}) : \mathbb{E}_{\mathbb{P}} [\tilde{z}] = \mu, \mathbb{E}_{\mathbb{P}} [\tilde{U}] = \Sigma, \right. \\ \left. \mathbb{P} \left(\begin{bmatrix} 1 & (\tilde{z} - \mu)^\top \\ (\tilde{z} - \mu) & \tilde{U} \end{bmatrix} \succcurlyeq 0 \right) = 1 \right\} \quad (12)$$

We then have $\mathbb{Q}^0 \in \Pi_{\tilde{z}} \mathcal{P} = \left\{ \mathbb{Q} \in \mathcal{P}_0(\mathbb{R}^P) : \mathbb{E}_{\mathbb{Q}} [\tilde{z}] = \mu, \mathbb{E}_{\mathbb{Q}} \left[(\tilde{z} - \mu) (\tilde{z} - \mu)^\top \right] \preceq \Sigma \right\}$.

Example 1.3 (The plus-function, high-order moment, ...)

$$\mathbb{E}_{\mathbb{P}} [|a_1^T z|] \leq \mu, \mathbb{E}_{\mathbb{P}} (a_2^T z)^2 \leq \delta, \mathbb{E}_{\mathbb{P}} [(a_3^T z)_+]^3 \leq \sigma, \dots$$

$$(1) \quad a_1^T z \leq \nu_1, -a_1^T z \leq \nu_1$$

$$(2) \quad \sqrt{(a_2^T z)^2 + \left(\frac{\nu_2 - 1}{2}\right)^2} \leq \frac{\nu_2 + 1}{2}$$

$$(3) \quad \mu_1 \geq 0, \mu_1 \geq a_3^T z,$$

$$[\mu_1^2 + \left(\frac{\mu_1 - 1}{2}\right)^2]^{1/2} \leq \frac{\mu_2 - 1}{2},$$

$$[\mu_2^2 + \left(\frac{\mu_1 - \nu_3}{2}\right)^2]^{1/2} \leq \frac{\nu_3 + \mu_1}{2}$$

Application in two-stage stochastic linear programming

$$\min_{x \in X} \left\{ c'x + \max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[Q(x, \tilde{z})] \right\}. \quad (13)$$

$$\begin{aligned} Q(x, \tilde{z}) &= \min_{y(\tilde{z})} && d'y(\tilde{z}) \\ &\text{s. t.} && A(\tilde{z})x + Dy(\tilde{z}) = b(\tilde{z}), \\ &&& y(\tilde{z}) \geq 0, \end{aligned}$$

$$\Rightarrow \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[Q(x, \tilde{z})] \leq t - c'x = w$$

A production planning problem. A company manager is considering the amount of steel to purchase (at \$58/lb) for producing wrenches and pliers in next month. The manufacturing process involves molding the tools on a molding machine and then assembling the tools on an assembly machine. Here is the technical data.

	Wrench	Plier	Total
Steel (lbs.)	1.5	1	x
Molding Machine (hours)	1	1	z_1
Assembly Machine (hours)	.3	.5	z_2
Contribution to Earnings (\$/1000 units)	130	100	

If z_1 and z_2 are fixed, then the problem is

$$\begin{aligned}
 \min \quad & 57x - 130w - 100p \\
 \text{s.t.} \quad & w + p \leq z_1 \text{ (Mold constraint)} \\
 & .3w + .5p \leq z_2 \text{ (Assembly constraint)} \\
 & 1.5w + p \leq x \text{ (Steel constraint)} \\
 & x, w, p \geq 0
 \end{aligned}$$

The Two-stage Stochastic Programming Formulation

Decision Variables: x : the steel to purchase now; w_i, p_i : production plan under scenario $i = 1, 2, 3, 4$.

Scenario	Assembly Hours	Molding Hours	Probability
1	8000	25000	.25
2	8000	21000	.25
3	10000	25000	.25
4	10000	21000	.25

We minimize the expected cost subject to “scenario constraints”.

$$\min \quad 58x - \sum_{i=1}^4 .25(130w_i + 100p_i)$$

$$\text{s.t.} \quad w_1 + p_1 \leq 25000$$

(Mold constraint for scenario 1)

$$.3w_1 + .5p_1 \leq 8000$$

(Assembly constraint for scenario 1)

$$-x + 1.5w_1 + p_1 \leq 0$$

(Steel constraint for scenario 1)

$$w_2 + p_2 \leq 21000$$

(Mold constraint for scenario 2)

$$.3w_2 + .5p_2 \leq 8000$$

(Assembly constraint for scenario 2)

$$\begin{aligned}
 -x + 1.5w_2 + p_2 &\leq 0 && \text{(Steel constraint for scenario 2)} \\
 w_3 + p_3 &\leq 25000 && \text{(Mold constraint for scenario 3)} \\
 .3w_3 + .5p_3 &\leq 10000 && \text{(Assembly constraint for scenario 3)} \\
 -x + 1.5w_3 + p_3 &\leq 0 && \text{(Steel constraint for scenario 3)} \\
 w_1 + p_1 &\leq 21000 && \text{(Mold constraint for scenario 4)} \\
 .3w_4 + .5p_4 &\leq 10000 && \text{(Assembly constraint for scenario 4)} \\
 -x + 1.5w_4 + p_4 &\leq 0 && \text{(Steel constraint for scenario 4)}
 \end{aligned}$$

$$x, w_i, p_i \geq 0, \quad i = 1, \dots, 4.$$

The solutions are as follows. $x = 27,250$, minimal expected cost = -802,833, and the production plans under various scenarios are as follows.

Scenario	w_i	p_i
1	12,500	8,500
2	12,500	8,500
3	8,056	15,167
4	12,500	8,500

The SOAD formulation. We have $X = \{x : x \geq 0\}$ and $\mathbb{P} \in \mathcal{P}$

$$\mathcal{P} := \left\{ \mathbb{P} : \begin{array}{l} \mathbb{E}_{\mathbb{P}}(\tilde{z}_j) = \mu_j, \quad j = 1, \dots, m, \\ \mathbb{E}_{\mathbb{P}}(\tilde{z}_j^2) \leq \eta_j, \quad j = 1, \dots, m, \\ \mathbb{P}\{\tilde{z} \in \Omega\} = 1. \end{array} \right\}$$

By Theorem 1.1 we solve the equivalent conic optimization problem:

$$\begin{array}{ll} \min_{x, v_0, v, V, u, Y, s, t, \lambda, \nu} & 58x + v_0 + 23000v_1 + 9000v_2 + 533 \times 10^9 V_1 + 82 \times 10^9 V_2 \\ \text{s. t.} & \sum_{i=1}^3 u_i - 130y_1^0 - 100y_2^0 + 25000\lambda_1 + 10000\lambda_2 + \lambda_3 - 21000\nu_1 \\ & -8000\nu_2 + \nu_3 - v_0 \leq 0, \\ & \left\| \begin{pmatrix} v_1 + 130y_1^1 + 100y_2^1 + \lambda_1 - \nu_1 \\ V_1 - u_1 \end{pmatrix} \right\| \leq V_1 + u_1, \\ & \left\| \begin{pmatrix} v_2 + 130y_1^2 + 100y_2^2 + \lambda_2 - \nu_2 \\ V_2 - u_2 \end{pmatrix} \right\| \leq V_2 + u_2, \\ & \left\| \begin{pmatrix} v_3 + 130y_1^3 + 100y_2^3 + \lambda_3 - \nu_3 \\ V_3 - u_3 \end{pmatrix} \right\| \leq V_3 + u_3, \\ & y_1^0 + y_2^0 + y_3^0 = 0, \quad .3y_1^0 + .5y_2^0 + y_4^0 = 0, \quad -x + 1.5y_1^0 + y_2^0 = 0, \\ & y_1^1 + y_2^1 + y_3^1 = 1, \quad .3y_1^1 + .5y_2^1 + y_4^1 = 0, \quad 1.5y_1^1 + y_2^1 = 0, \\ & y_1^2 + y_2^2 + y_3^2 = 0, \quad .3y_1^2 + .5y_2^2 + y_4^2 = 1, \quad 1.5y_1^2 + y_2^2 = 0, \\ & y_1^3 + y_2^3 + y_3^3 = 0, \quad .3y_1^3 + .5y_2^3 + y_4^3 = 0, \quad 1.5y_1^3 + y_2^3 = 1 \end{array}$$

$$\begin{aligned}
y_1^1 + y_2^1 + y_3^1 &= 1, & 0.3y_1^1 + 0.5y_2^1 + y_4^1 &= 0, & 1.5y_1^1 + y_2^1 &= 0, \\
y_1^2 + y_2^2 + y_3^2 &= 0, & 0.3y_1^2 + 0.5y_2^2 + y_4^2 &= 1, & 1.5y_1^2 + y_2^2 &= 0, \\
y_1^3 + y_2^3 + y_3^3 &= 0, & 0.3y_1^3 + 0.5y_2^3 + y_4^3 &= 0, & 1.5y_1^3 + y_2^3 &= 1, \\
-21000s_1^1 - 8000s_2^1 + s_3^1 + 25000t_1^1 + 10000t_2^1 + t_3^1 - y_1^0 &\leq 0, \\
-21000s_1^2 - 8000s_2^2 + s_3^2 + 25000t_1^2 + 10000t_2^2 + t_3^2 - y_2^0 &\leq 0, \\
-21000s_1^3 - 8000s_2^3 + s_3^3 + 25000t_1^3 + 10000t_2^3 + t_3^3 - y_3^0 &\leq 0, \\
-21000s_1^4 - 8000s_2^4 + s_3^4 + 25000t_1^4 + 10000t_2^4 + t_3^4 - y_4^0 &\leq 0,
\end{aligned}$$

$$\begin{aligned}
s_1^1 - t_1^1 - y_1^1 &\leq 0, & s_2^1 - t_2^1 - y_1^2 &\leq 0, & s_3^1 - t_3^1 - y_1^3 &\leq 0, \\
s_1^2 - t_1^2 - y_2^1 &\leq 0, & s_2^2 - t_2^2 - y_2^2 &\leq 0, & s_3^2 - t_3^2 - y_2^3 &\leq 0, \\
s_1^3 - t_1^3 - y_3^1 &\leq 0, & s_2^3 - t_2^3 - y_3^2 &\leq 0, & s_3^3 - t_3^3 - y_3^3 &\leq 0, \\
s_1^4 - t_1^4 - y_4^1 &\leq 0, & s_2^4 - t_2^4 - y_4^2 &\leq 0, & s_3^4 - t_3^4 - y_4^3 &\leq 0, \\
x &\geq 0, & V, u, \lambda, \nu &\geq 0, & s^k, t^k &\geq 0, & k = 1, 2, 3, 4.
\end{aligned}$$

The numerical results show that $x = 31,500$, max-worst cost = -940,770.

(Gao, S, Wu) Based various levels of knowledge about moments, different steel purchasing levels are calculated respectively and the comparison is shown in the table below.

Number of moments known	Steel purchased
1	30500
2	30500
3	30500
4	30500
5	27861
6	17876
7	17876

It is interestingly noted adding one extra moment information may have no value in the sense of making SOAD decision . When cost of evaluating uncertainty moments is high, dealing with low level of knowledge can be sufficient to worst-case decision making.

2 The chance constraint: A non-convex case

$$\mathbb{P}(a(\tilde{z})^T x \geq b(\tilde{z})) \geq 1 - \epsilon \quad (14)$$

For mathematicians,

$$\mathbb{E}(\mathbf{1}_{a(\tilde{z})^T x \geq b(\tilde{z})}) \geq 1 - \epsilon \quad (15)$$

It is easy to see that $\mathbf{1}_{a(\tilde{z})^T x \geq b(\tilde{z})}$ is not bilinear in (x, \tilde{z}) so the convex SOAD framework does not apply.

For financial experts,

$$\mathbb{P}[a(\tilde{z})^T x \geq b(\tilde{z})] \geq 1 - \epsilon \iff \text{VaR}_{1-\epsilon}(b(\tilde{z}) - a(\tilde{z})^T x) \leq 0.$$

Define (Rockafellar and Uryasev)

$$\text{CVaR}_{1-\epsilon}(X) = \min_{\beta} \left\{ \beta + \frac{1}{\epsilon} \mathbb{E}[(X - \beta)_+] \right\}, \quad (16)$$

where X is a single random variable. If $X = v(x, \tilde{z})$ is bilinear in (x, \tilde{z}) (or convex in x), then $\text{CVaR}(v(x, \tilde{z}))$ is convex in x .

It can be shown that

$$\text{VaR}_{1-\epsilon}(X) \leq \text{CVaR}_{1-\epsilon}(X).$$

Thus, CVaR is a convex upper bound of VaR. In fact, it is best c.u. of it (Nemirovski and Shapiro).

By bounding $\mathbb{E}((y_0 + \mathbf{y}'\tilde{\mathbf{z}})_+)$ one can reduce the case to LP, SOCP, or SDP. Our computational test shows that even the LP bounding method is satisfactory for a resource allocation problem.

See details in *Operations Research* 58 (2010) 470-485.

Thank You!