STOCHASTIC PROGRAMMING

December 21, 2023

Today: last lecture (probabilistic constraints)

Next step: Exam, January 15 10am - 12pm
Allowed: 1 "cheatsheet" of notes
A4, two-sided

Exercise from session 4

Mean deviation model

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \mathbb{E}_\xi \left[\, | \xi^T x | \, \right] \quad \text{s.t.} \quad \sum_{i=1}^n x_i = 1, \quad x \geq 0$$

x : portfolio allocation

ξ : random vector of returns

a) Reformulation as a linear two-stage stochastic program

↳ We first introduce a second-stage optimization problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \mathbb{E}_\xi \left[Q(x, \xi) \right] \quad \text{s.t.} \quad \sum_{i=1}^n x_i = 1 \; , \; x \geq 0$$

where $Q(x, \xi) = \underbrace{| \xi^T x |}_{\text{Not a linear function}}$

For any real number $a \in \mathbb{R}$, we can write a as $a^+ - a^-$ where $a^+ = \max(a, 0) \geq 0$ and $a^- = \max(-a, 0) \geq 0$

$$a = a^+ - a^- \implies |a| = a^+ + a^-$$

Hence we can rewrite $\xi^T x$ as $y^+ - y^-$ where $y^+ \geq 0$ and $y^- \geq 0$. In that case, $|\xi^T x| = y^+ + y^-$

$$Q(x, \xi) = |\xi^T x| = \left| \begin{array}{ll} \underset{\substack{y^+ \in \mathbb{R} \\ y^- \in \mathbb{R}}}{\min} \; y^+ + y^- & \text{s.t.} \; y^+ - y^- = \xi^T x \\ & \quad y^+ \geq 0 \\ & \quad y^- \geq 0 \end{array} \right.$$

Linear program in y^+, y^- →

→ linear constraint w.r.t. x

$$Q(x,\xi) = |\xi^T x| = \min_{y\in\mathbb{R}} y \quad \text{s.t.} \quad y = |\xi^T x|$$

b) Scenario approximation

$\xi_1, \dots, \xi_K$ with $\mathbb{P}(\xi = \xi_k) = p_k \quad \forall k = 1..K$

$\forall k = 1..K$, the second-stage problem is

$$\begin{aligned} \underset{\substack{y_k^+ \in\mathbb{R} \\ y_k^- \in\mathbb{R}}}{\text{minimize}} \quad & y_k^+ + y_k^- \quad \text{s.t.} \quad & y_k^+ - y_k^- = \xi_k^T x \\ & & y_k^+ \geq 0 \\ & & y_k^- \geq 0 \end{aligned}$$

$$\mathbb{E}_\xi[Q(x,\xi)] \longrightarrow \sum_{k=1}^{K} p_k\, Q(x,\xi_k)$$

First scenario formulation

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad \sum_{k=1}^{K} p_k\, Q(x,\xi_k) \quad \text{s.t.} \quad \sum_{i=1}^{n} x_i = 1,\ x\geq 0$$

where $$Q(x,\xi_k) = \min_{\substack{y_k^+ \in\mathbb{R} \\ y_k^- \in\mathbb{R}}} \left\{ y_k^+ + y_k^- \;\middle|\; y_k^+ - y_k^- = \xi_k^T x,\ y_k^+ \geq 0,\ y_k^- \geq 0 \right\}$$

Second reformulation as a single linear program $\quad p = \begin{bmatrix} p_1 \\ \vdots \\ p_K \end{bmatrix} \in \mathbb{R}^K$

$$\begin{aligned} \underset{\substack{x\in\mathbb{R}^n \\ y^+\in\mathbb{R}^K \\ y^-\in\mathbb{R}^K}}{\text{minimize}} \quad & p^T y^+ + p^T y^- \\ \text{s.t.} \quad & y_k^+ - y_k^- = \xi_k^T x \quad \forall k = 1..K \\ & \sum_{i=1}^{n} x_i = 1 \end{aligned}$$

$$A\begin{bmatrix} x \\ y^+ \\ y^- \end{bmatrix} = b \qquad b = \begin{array}{c} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow K \\ \\ \end{array}$$

$x \geq 0$
$y^+ \geq 0$
$y^- \geq 0$

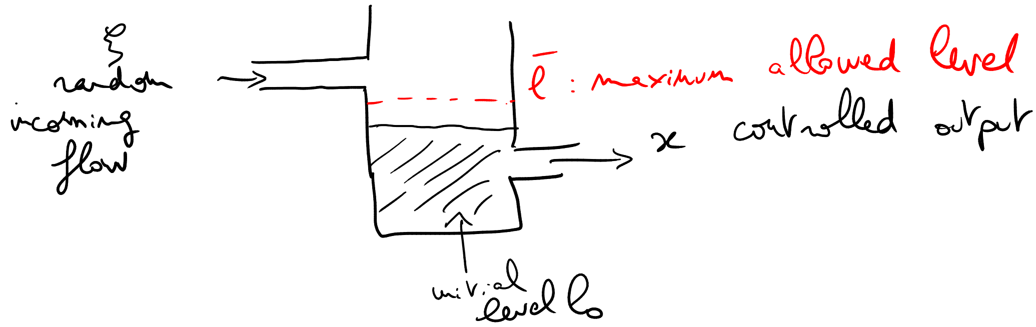
$$A = \begin{bmatrix} \xi_1^T & -1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \xi_2^T & 0 & -1 & 0 & \cdots 0 & 0 & 1 & 0 \cdots & 0 \\ & & & \vdots & & & & & \\ \xi_K^T & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 & 1 \\ 1 \cdots 1 & 0 & & \cdots & & & & & 0 \end{bmatrix} \begin{array}{l} \to \xi_1^T x - y_1^+ + y_1^- = 0 \\ \\ \\ \to \xi_K^T x - y_K^+ + y_K^- = 0 \\ \to \sum_{i=1}^{m} x_i = 1 \end{array}$$

$$A = \begin{bmatrix} \boxed{\Psi}_K & -I_K & I_K \\ 1 \cdots 1 & 0 \cdots & 0 \end{bmatrix}$$

$$\boxed{\Psi}_K = \begin{bmatrix} \xi_1^T \\ \vdots \\ \xi_K^T \end{bmatrix} \in \mathbb{R}^{K \times m}$$

PROBABILISTIC CONSTRAINTS

① Motivating example : Level constraints in a reservoir



→ Several time steps $t=1 \dots T$

. At each step, an amount ξ_i is coming in and we remove a quantity x_i

. Constraint: maximum level should never be exceeded after every time step

T constraints
$$\begin{cases} \forall t \in \{1, \dots, T\}, \\ \underbrace{\ell_0 + \xi_1 + \dots + \xi_t - x_1 - \dots - x_t}_{\text{level after step } t} \leq \bar{\ell} \end{cases}$$

$$\iff \quad L\xi \leq Lx + \ell$$

$u \leq v \iff u_i \leq v_i \ \forall i = 1..T$, $u \in \mathbb{R}^T$

$$L = \begin{bmatrix} 1 & & 0 \\ \vdots & \ddots & \\ 1 & \cdots & 1 \end{bmatrix} \qquad \ell = \begin{bmatrix} \bar{\ell} - \ell_0 \\ \vdots \\ \bar{\ell} - \ell_0 \end{bmatrix} \in \mathbb{R}^T$$

↳ Guaranteeing that $L\xi \leq Lx + \ell$ is satisfied deterministically (for all realizations of ξ) is often considered as too expensive

⇒ The constraint is replaced by a probabilistic constraint

$$\mathbb{P}_\xi(L\xi \leq Lx + \ell) \geq p$$, where p is close to 1

$p = 0.95$ or $p = 0.99$

1 Definition and examples

A general form for probabilistic constraints is

$$\mathbb{P}(\, g(x,\xi) \geq 0 \,) \geq p$$

$x \in \mathbb{R}^n$: optimization variable

$\xi \in \mathbb{R}^m$: uncertainty

$g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^\ell$

$p \in [0,1]$

$p=0$: always true
(deterministic constraint, no constraints on x)

$p=1$: almost sure constraint
$\Rightarrow \mathbb{P}(g(x,\xi) \geq 0) = 1$

↳ We distinguish between two types of constraints

- Joint probabilistic constraints

$$\mathbb{P}(\, g_i(x,\xi) \geq 0 \quad \forall i = 1..\ell \,) \geq p$$

- Individual probabilistic constraints

$$\mathbb{P}(\, g_i(x,\xi) \geq 0 \,) \geq p \qquad \forall i = 1..\ell$$

↳ Special cases

- Linear probabilistic constraint

Joint constraint →

$$\mathbb{P}(\, \underbrace{A(x)\xi}_{\text{Linear in } \xi} \leq b(x) \,) \geq p$$

$A(x) \in \mathbb{R}^{\ell \times m}$

$b(x) \in \mathbb{R}^\ell$

ℓ : number of rows of $A(x)$

- When $A(x) = A \;\; \forall x \in \mathbb{R}^n$, this constraint is called a separable linear constraint

$$\underbrace{A}_{\text{does not depend on } x}\xi \leq b(x)$$

Remark: Even for linear constraints, hard to understand/predict the shape/landscape of the constraint function

$$x \longmapsto \mathbb{P}(A(x)\xi \leq b(x))$$

Ex (R. Henrion, 2014)

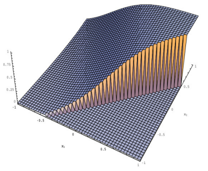
$$x \in \mathbb{R}^2 \mapsto \mathbb{P}(L\xi + Mx \geq b)$$

$\xi \sim \mathcal{N}(0,1)$ in $\mathbb{R}$, M fixed to $\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$

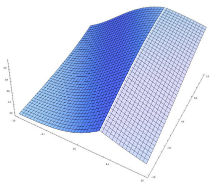
Two variants: 1) $L = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ -0.5 \end{bmatrix}$

$$\mathbb{P}\left(\begin{array}{l} -\xi + 2x_1 + x_2 \geq 0 \\ \text{and} \\ -x_1 + x_2 \geq -0.5 \end{array} \right)$$

$(M|L|b) = \left(\begin{array}{cc|c|c} 2 & 1 & -1 & 0 \\ -1 & 1 & 0 & -0.5 \end{array} \right)$



$(M|L|b) = \left(\begin{array}{cc|c|c} 2 & 1 & -1 & 0 \\ -1 & 1 & -1 & -1 \end{array} \right)$



2) $L = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

$$\mathbb{P}(-\xi + 2x_1 + x_2 \geq 0 \text{ and } -\xi - x_1 + x_2 \geq -1)$$

② Formulations of probabilistic constraints and feasible sets

↳ In general, a probabilistic constraint cannot be reformulated as a deterministic constraint

⇒ Still, possible for <u>some</u> classes of probabilistic constraints.

a) Individual chance constraints of the form

$$g(x, \xi) = h(x) - \xi \qquad \text{where } h: \mathbb{R}^n \to \mathbb{R}, \ \xi \in \mathbb{R}$$

$$\mathbb{P}(g(x,\xi) \geq 0) = \underbrace{\mathbb{P}(\xi \leq h(x))}_{\text{cumulative distribution function of } \xi \text{ evaluated at } h(x)}$$

The cumulative distribution can be inverted for many standard distributions (Gaussian, exponential, Beta, ...)

$\Rightarrow$ Not necessarily a closed-form expression but good numerical approximations

$\Rightarrow$ We say that such constraints $\mathbb{P}(\xi \leq h(x)) \geq p$ are <u>explicit</u> chance constraints, in that

$$\mathbb{P}(\xi \leq h(x)) \geq p \iff \overbrace{h(x) \geq q(p)}^{\text{Deterministic constraint on } x}, \quad \text{where}$$

$q(p)$ is the p-quantile of the distribution of ξ, i.e. $\mathbb{P}(\xi \leq q(p)) = p$

<u>Remark</u>:

- This approach extends to multiple individual constraints

$$\mathbb{P}(\xi \leq h_i(x)) \geq p \qquad \forall i = 1..\ell$$

$$\iff \quad h_i(x) \geq q(p) \qquad \forall i = 1..\ell$$

- But it does not imply that the result holds for joint constraints

$$\mathbb{P}(\underbrace{\xi \leq h_i(x) \ \forall i = 1..\ell}_{\text{events are not independent because they apply to the same } \xi}) \geq p$$

$$\neq$$

$$\mathbb{P}(\underbrace{\xi \leq h_i(x)}_{\text{events are independent}}) \geq p \quad \forall i = 1..\ell$$

$\rightarrow$ One can apply the reasoning to $\mathbb{P}(\xi \leq h_i(x) \ \forall i = 1..\ell) \geq p$ directly but cannot conclude from the individual constraint case.

More generally, when ξ is Gaussian, can use the cumulative distribution to evaluate the constraint

$\Rightarrow$ when $g(x,\xi) = A(x)\,\xi - b(x)$, the evaluation of the probabilistic constraint (and possibly its derivative, which may needed in optimization routines) reduces to evaluating the CDF (cumulative distribution function)

$\hookrightarrow$ For optimization, we also want to find conditions that certify that a feasible set defined by probabilistic constraints is convex.

$\hookrightarrow$ $\mathcal{F}$ is convex in $\mathbb{R}^n$ if $\forall (x,y) \in \mathcal{F}^2$, $\forall \alpha \in [0,1]$, $\alpha x + (1-\alpha) y \in \mathcal{F}$

$\hookrightarrow$ Not straightforward in the general case, but guarantees exist for linear constraints

Theorem: Let $A \in \mathbb{R}^{l \times m}$, $p \in [0,1]$, $h : \mathbb{R}^n \to \mathbb{R}^l$, $h = [h_j]$

Consider the set

$$\mathcal{F} = \{ x \in \mathbb{R}^n \mid \mathbb{P}(A\xi \leq h(x)) \geq p \}$$

(joint constraint, linear in ξ, separable)

$\mathcal{F}$ is convex if the cumulative distribution of ξ is log-concave and every h_j is concave

$\log(\;)$ is concave

$h : \mathbb{R}^n \to \mathbb{R}$ concave if

$$h(\alpha x + (1-\alpha) y) \geq \alpha h(x) + (1-\alpha) h(y)$$

$\forall x, y$, $\forall \alpha \in [0,1]$

$e^{-(\;)}$

Convex feasible set $\Rightarrow$
- Optimality conditions
- Uniquely defined projections

$\Rightarrow$ Easy to optimize a convex set

(Deterministic) Convex programming algorithms and software are used in stochastic programming as tools, akin to linear programming solvers.

Ex) $\xi \sim N(\mu, \Sigma)$ $\quad \mu \in \mathbb{R}^m$, $\quad \Sigma \in \mathbb{R}^{m \times m}$

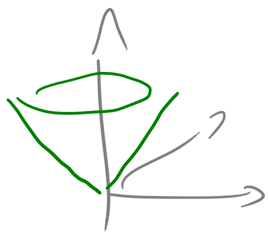
$\Sigma = \Sigma^T \succeq 0$

$$\{x \in \mathbb{R}^n \mid \mathbb{P}(a \le Ax + B\xi \le b) \ge p\} \text{ is convex}$$

$\forall a \le b$, $\forall A, B$

$X \succeq 0 \qquad v^T X v \ge 0 \ \ \forall v \in \mathbb{R}^n$

$$\|Ax + b\|^2 \le (c^T x + d)^2$$



$$x_2 \ge x_1^2$$

③ Scenarios and chance constraints

↳ When the distribution of ξ is not known, we use scenarios to approximate the probabilistic constraint

$\xi_1, \dots, \xi_K$ (suppose that they have the same probability)

↳ $\mathbb{P}(g(x, \xi) \ge 0)$

is approximated by

cardinality

$$\frac{1}{K} \left| \underbrace{\{k \in \{1, \dots, K\} \mid g(x, \xi_k) \ge 0\}}_{\subseteq \{1, \dots, K\}} \right|$$

or equivalently

$$\frac{1}{K}\sum_{i=1}^{K} \mathbb{1}(g(x,\xi_k)\geq 0)$$

$$\mathbb{1}(g(x,\xi_k)\geq 0) = \begin{cases} 1 & \text{if } g(x,\xi_k)\geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Generalization $\sum_{i=1}^{K} p_k \mathbb{1}(g(x,\xi_k)\geq 0)$ $\quad \sum_{k=1}^{K} p_k = 1 \quad p_k \geq 0$

$\hookrightarrow$ $\mathbb{P}(g(x,\xi)\geq 0)\geq p$ becomes

$$\frac{1}{K}\sum_{k=1}^{K} \mathbb{1}(g(x,\xi_k)\geq 0) \geq p$$

which is equivalent to

$$\frac{1}{K}\sum_{k=1}^{K}(1-s_k)\geq p \qquad s_k\in\{0,1\}$$

$$g(x,\xi_k)\geq \tau_k s_k \qquad \forall k=1..K$$

$\tau_k \leq \min_x g(x,\xi_k)$ fixed value

$s_k = 1$: No restriction
$s_k = 0$: $g(x,\xi_k)\geq 0$ necessarily

Ex) $\quad g(x,\xi) = a^T x - \xi \qquad x\in\mathcal{X} \qquad a^T x \geq 0 \;\; \forall x\in\mathcal{X}$

$$\mathbb{P}(a^T x - \xi \geq 0)\geq p$$

can be estimated by

$$\frac{1}{K}\sum_{k=1}^{K} s_k \geq p \quad \text{where} \quad s_k\in\{0,1\}$$

$$a^T x - \xi_k \geq -\xi_k s_k \qquad \forall k=1..K$$

$\Rightarrow$ Adding variables to represent satisfaction of $g_i(x,\xi_k)\geq 0$ gives rise to a mixed-integer optimization problem (harder to solve than continuous optimization problem but still computational tractable for a moderate number of scenarios ≤ 100 with a convex problem and even beyond for linear objective and constraints)

↳ <u>Alternate approach</u>

Replace the probabilistic constraints by the scenario constraints

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x) \quad \text{s.t.} \quad \mathbb{P}\left(g_i(x,\xi)\geq 0 \ \forall i=1..\ell\right)\geq p$$

$$\Big\Downarrow \quad \xi_1, \dots, \xi_K \ \text{scenarios generated randomly}$$

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x) \quad \text{s.t.} \quad g_i(x,\xi_k)\geq 0 \quad \begin{array}{l}\forall i=1..\ell\\ \forall k=1..K.\end{array}$$

<u>Th)</u> (2005)

Suppose that $K > \dfrac{n}{\varepsilon\beta} - 1$ for some $\varepsilon \in (0, 1-p]$ and $\beta \in (0,1)$

Consider $x_K^* \in \underset{x\in\mathbb{R}^n}{\operatorname{argmin}} f(x)$ s.t. $g_i(x,\xi_k)\geq 0$ $\begin{array}{l}i=1..\ell\\ k=1..K\end{array}$

Then $\mathbb{P}\Big(\underline{\mathbb{P}(g_i(x_K^*,\xi)\geq 0)\geq p}\Big) \geq 1-\beta$

Probability over scenarios/samples $\xi_1, \dots, \xi_K$

Probability w.r.t. ξ original constraint

↳ Modern chance-constrained programming relies on statistics to provide guarantees for scenario formulations.

Ex) (Ahmed & Luedtke 2008)

$$v^* := \underset{x \in \mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.} \quad \mathbb{P}(g(x,\xi) \geq 0) \geq p$$

Scenario version

$$v^*_{K,q} = \underset{x \in \mathbb{R}^n}{\text{argmin}}\ f(x) \quad \text{s.t.} \quad \frac{1}{K}\sum_{h=1}^{K} \mathbb{1}(g(x,\xi_h) \geq 0) \geq q$$

$q > p$

Using concentration inequalities, one can show that

$$\mathbb{P}(v^*_{K,q} \leq v^*) \geq 1 - \exp(-2K(q-p)^2)$$

$$q = 1 \Rightarrow g(x, \xi_k) \geq 0 \ \forall k = 1..K$$

↳ This approach gives a lower bound on the optimum whereas the previous approach (using scenarios directly) gives a feasible point, hence an upper bound on v^*

Overall, solving chance-constrained problems requires:

→ An approach to generate scenarios (ideally knowledge about the distribution of ξ)

→ Good deterministic solvers for the class of problems of interest (convex programming, mixed-integer convex programming, nonlinear nonconvex programming, etc)

Bibliography

→ Optimization methods in finance, 2018 (Peña, Cornuejols, Tütüncü)

→ Introduction to stochastic programming 2011 (Birge, Louveaux)

→ Lectures on stochastic programming, 2021 (Shapiro, Ruszczyński, Dentcheva)

→ A tutorial on stochastic programming 2007 (Shapiro, Philpott)

→ stoprog.org

→ Vincent Leclère (ENPC)