Power System Analysis

Chapter 13 Stochastic optimal power flow

Steven Low Caltech (Feb 25, 2025)

Stochastic OPF

Consider

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad h(x, \zeta) \leq 0$$

where ζ is a parameter, e.g., admittance matrix, renewable generations, forecast loads

In many power system applications some of these parameters are uncertain, giving rise to stochastic OPF

Brief introduction to theory of stochastic optimization

- Most stochastic optimization problems are intractable (e.g., nonconvex, nonsmooth)
- Explains 4 main ideas to deal with uncertainty
- Focuses on convex reformulations and structural properties

Stochastic OPF

4 main ideas

Choose optimal x^* s.t.

- Robust opt: x^* satisfies constraints for all ζ in an uncertainty set Z
- Chance constrained opt: x^* satisfies constraints with high probability
- Scenario opt: x^* satisfies constraints for K random samples of $\zeta \in Z$
- Two-stage opt: 2nd-stage decision $y(x^*, \zeta)$ adapts to realized parameter ζ, given 1st-stage decision x^*

Many methods are combinations of these 4 ideas, e.g.

- Distributional robust opt: robust + chance constrained
- Adaptive robust opt: two-stage + robust (as opposed to expected) 2nd-stage cost
- Adaptive robust affine control: two-stage + robust (or avg) + affine policy

Outline

1. Robust optimization
2. Chance constrained optimization
3. Convex scenario optimization
4. Stochastic optimization with recourse

Outline

1. Robust optimization
    - General formulation
    - Robust linear program
    - Robust second-order cone program
    - Robust semidefinite program
    - Proofs
2. Chance constrained optimization
3. Convex scenario optimization
4. Stochastic optimization with recourse

General formulation

Consider

$$\min_{x \in \mathbb{R}^n} \quad f(x) \quad \text{s.t.} \quad h(x, \zeta) \;\leq\; 0, \quad \forall \zeta \in Z(x)$$

- $f(x)$: cost function is assumed certain wlog
- ζ : uncertain parameter
- $h(x, \zeta)$: uncertain inequality constraint
- $Z(x)$: uncertainty set that can depend on optimization variable x

Interpretation: Choose an optimal x^* that satisfies the inequality constraint $h(x^*, \zeta) \leq 0$ for all possible uncertainty realization $\zeta \in Z(x^*)$

General formulation

Consider

$$\min_{x \in \mathbb{R}^n} \quad f(x) \quad \text{s.t.} \quad h(x, \zeta) \ \leq \ 0, \quad \forall \zeta \in Z(x)$$

- Semi-infinite program: finite #optimization variables $x \in \mathbb{R}^n$, possibly infinite #constraints
- Generally intractable
- For special cases of uncertainty set $Z(x)$, robust program has finite convex reformulation which is tractable
- e.g. robust LP, robust SOCP, robust SDP

General formulation

Example

- PV panel with uncertain real power generation $\zeta_t \in Z_t \subseteq \mathbb{R}_+$ and controllable reactive power $q_t \in [q^{\min}, q^{\max}]$
- PV panel is connected to battery through a line with series admittance $y \in \mathbb{C}$
- DC discharging power $d_t \in [d^{\min}, d^{\max}]$ is controllable as long as its SoC $b := (b_1, \ldots, b_T)$ satisfies $b_t \in [0, B]$
- Voltages at buses 1 and 2 are $v_{1t} = |v_{1t}| e^{i\theta_{1t}}, \; v_{2t} = |v_{2t}| e^{i\theta_{2t}}$. Let $v_t := (v_{1t}, v_{2t})$

Goal: control (q_t, d_t) within control limits at time t to min cost, subject to SoC $b_t \in [0, B]$ and voltage limits $|v_{it}| \in [v^{\min}, v^{\max}]$ for $t = 1, \ldots, T$

General formulation

Example

Let $x := (q, d) \in \mathbb{R}^{2T}, v := (v_1, \ldots, v_T), b := (b_1, \ldots, b_T), \zeta := (\zeta_1, \ldots, \zeta_T)$

Robust scheduling problem is:

$$\min_x \ f(x) \quad \text{s.t.} \quad g(x, v, b, \zeta) = 0, \ h(x, v, b, \zeta) \leq 0, \ \forall \zeta \in Z_1 \times \cdots \times Z_T$$

where $g(x, v, b, \zeta) = 0$ are power equation and battery state process

$$\zeta_t + iq_t = \ y^{\mathsf{H}} \left(|v_{1t}|^2 - v_{1t} v_{2t}^{\mathsf{H}} \right), \qquad d_t + i0 = \ y^{\mathsf{H}} \left(|v_{2t}|^2 - v_{2t} v_{1t}^{\mathsf{H}} \right), \qquad b_{t+1} = \ b_t - d_t$$

and $h_t(x, t, b, \zeta) \leq 0$ are voltage and battery limits

$$v^{\min} \leq |v_{it}| \leq v^{\max}, \ \ i = 1,2, \qquad 0 \leq b_t \leq B$$

uncertain equality constraints need to be interpreted appropriately and eliminated

General formulation

Example

Given control decisions $x_t := (q_t, d_t)$ and uncertain parameter ζ_t , voltage v_t takes value in

$$V_t(x) := \{v_t \in \mathbb{C}^2 : v_t \text{ satisfies power flow equation, } \zeta_t \in Z_t\}$$

To eliminate battery, write b_t as

$$b_t \;=\; b_0 - \sum_{s<t} d_s, \qquad t = 1,\ldots,T$$

Then robust scheduling problem is:

$$\min_x \; f(x) \quad \text{s.t.} \quad v^{\min} \le |v_{it}| \le v^{\max},\ i = 1,2,\ \forall v_t \in V_t(x),\ t = 1,\ldots,T$$

$$0 \le b_0 - \sum_{s<t} d_s \le B,\ t = 1,\ldots,T$$

The original uncertainty set Z_t is embedded into the x-dependent uncertainty set $V_t(x)$

General formulation

Tractability

Consider

$$\min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s.t.} \quad h(x, \zeta) \;\leq\; 0, \quad \forall \zeta \in Z(x)$$

Equivalent bi-level formulation

$$\min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s.t.} \quad \sup_{\zeta \in Z(x)} \; h(x, \zeta) \;\leq\; 0 \tag{1}$$

Assuming f is convex, tractability of (1) boils down to whether the following subproblem is tractable:

$$\bar{h}(x) := \sup_{\zeta \in Z(x)} \; h(x, \zeta)$$

Derivation strategy

3 common strategies to derive finite convex reformulation of robust optimizations:

1. Solve $\bar{h}(x)$ analytically in close form and $\bar{h}(x) \leq 0$ is convex in x — robust LP
2. Replace $\bar{h}(x) \leq 0$ by strong duality $d(y) \leq 0$ and KKT condition such that y is optimal for the dual of subproblem $\sup_{\zeta \in Z(x)} h(x, \zeta)$, i.e., y satisfies dual feasibility and stationary
   (a) Need Slater theorem ($\bar{h}(x)$ is finite, convexity and Slater condition) to guarantee strong duality and existence of dual opt y — robust LP
   (b) ζ is eliminated because (i) $h(x, \zeta)$ is affine in ζ and therefore $\nabla_\zeta L(\zeta, y) = 0$ does not contain ζ; and (ii) in strong duality and stationarity imply complementary slackness (which therefore can be omitted)

Derivation strategy

3. When the semi-infinite constraint takes the form $h_0(x) + h(x, \zeta) \in K$ for all $\zeta \in Z$ where K is a closed convex cone, such as $K_{\text{soc}} \subseteq \mathbb{R}^n$ or $K_{\text{sdp}} \subseteq \mathbb{S}^n$, it can be reformulated as a set of linear matrix inequalities (LMIs) using the S-lemma. The resulting problem is an SDP

robust SOCP, robust SDP

Robust linear program

Consider

$$\min_{x \in \mathbb{R}^n} \; c^\top x \;\; \text{s.t.} \;\; a^\top x \le b, \; \forall [a^\top \; b] \in \left\{ \left[a_0^\top \; b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top \; b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

$(a_0, b_0) \in \mathbb{R}^{n+1}$ are nominal parameters; $\sum_l \zeta_l \left[a_l^\top \; b_l\right]$ are perturbations, with given $\left[a_l^\top \; b_l\right]$

Constraints are equivalent to

$$\bar{h}(x) := \max_{\zeta \in Z} \sum_{l=1}^{k} \zeta_l (a_l^\top x - b_l) \;\; \le -\, (a_0^\top x - b_0)$$

$\zeta := (\zeta_1, \ldots, \zeta_k)$ takes value in uncertainty set Z

This is general and allows each entry of a, b to vary independently (with $k = n + 1$)

Robust linear program

Consider

$$\min_{x\in\mathbb{R}^n} c^\top x \quad \text{s.t.} \quad a^\top x \le b,\ \forall [a^\top\ b] \in \left\{ \left[a_0^\top\ b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top\ b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

Theorem

1. Linear uncertainty $Z := \{\zeta \in \mathbb{R}^k : \|\zeta\|_\infty \le 1\}$: (1) is equivalent to LP: strategy 1

$$\min_{(x,y)\in\mathbb{R}^{n+k}} c^\top x \quad \text{s.t.} \quad a_0^\top x + \sum_l y_l \le b_0, \ -y_l \le a_l^\top x - b_l \le y_l, \ \ l = 1,\ldots,k$$

2. SOC uncertainty $Z := \{\zeta \in \mathbb{R}^k : \|\zeta\|_2 \le r\}$: (1) is equivalent to SOCP: strategy 1

$$\min_{x\in\mathbb{R}^n} c^\top x \quad \text{s.t.} \quad r\sqrt{\sum_l \left(a_l^\top x - b_l\right)^2} \le -a_0^\top x + b_0$$

Robust linear program

Consider

$$\min_{x\in\mathbb{R}^n} \ c^\top x \ \text{ s.t. } \ a^\top x \le b, \ \forall [a^\top \ b] \in \left\{ \left[a_0^\top \ b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top \ b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

Theorem

3. Conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$ where $K \subseteq \mathbb{R}^m$ is a closed convex pointed cone with nonempty interior.

Example: $Z := \{\zeta \in \mathbb{R}^k : \zeta \in K\}$

Conic uncertainty of part 3 is very general and includes parts 1 and 2 as special cases

Robust linear program

Consider

$$\min_{x \in \mathbb{R}^n} c^\top x \text{ s.t. } a^\top x \leq b, \forall [a^\top \ b] \in \left\{ \left[a_0^\top \ b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top \ b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

Theorem

3. Conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$ where $K \subseteq \mathbb{R}^m$ is a closed convex pointed cone with nonempty interior.

Example: $Z := \{\zeta \in \mathbb{R}^k : \zeta \in K\}$. Then (1) is equivalent to conic program: $\min_{(x,y) \in \mathbb{R}^{n+m}} c^\top x$ s.t.

$$a_0^\top x \leq b_0, \quad a_l^\top x + y_l = b_l, \ y \in K^*, \ l = 1, \ldots, k$$

Robust linear program

Consider

$$\min_{x \in \mathbb{R}^n} c^\top x \;\text{ s.t. }\; a^\top x \le b,\ \forall [a^\top\ b] \in \left\{ \left[a_0^\top\ b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top\ b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

Theorem

3. Conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$ where $K \subseteq \mathbb{R}^m$ is a closed convex pointed cone with nonempty interior.

The subproblem in the bi-level formulation is

$$\bar{h}(x) := \max_{\zeta \in Z} \sum_{l=1}^{k} \zeta_l (a_l^\top x - b_l) = \max_{(\zeta, u) \in \mathbb{R}^{k+p}} (s(x))^\top \zeta \;\text{ s.t. }\; \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} \zeta \\ u \end{bmatrix} + d \in K$$

this subproblem will be replaced by strong duality and KKT condition for $\bar{h}(x)$

Robust linear program

Consider

$$\min_{x\in\mathbb{R}^n} \ c^\top x \ \text{ s.t. } \ a^\top x \le b, \ \forall [a^\top \ b] \in \left\{ \left[a_0^\top \ b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top \ b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

Theorem

3. Conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$ where $K \subseteq \mathbb{R}^m$ is a closed convex pointed cone with nonempty interior. Suppose Z is nonempty and
   - Slater condition: either K is polyhedral cone or $\exists(\bar{\zeta}, \bar{u}) \in \mathbb{R}^{k+p}$ s.t. $P\bar{\zeta} + Q\bar{u} + d \ \in \ \text{ri}(K)$
   - For each x, $\max_{\zeta\in Z} \sum_l \zeta_l \left(a_l^\top x - b_l\right)$ is finite

   Then (1) is equivalent to conic program: $\min_{(x,y)\in\mathbb{R}^{n+m}} \ c^\top x$ s.t.

   $$a_0^\top x + d^\top y \le b_0, \quad y \in K^*, \quad Q^\top y = 0, \ a_l^\top x + \left(P^\top y\right)_l = b_l, \ l = 1,\ldots,k$$

   strong duality — dual feasibility — stationarity

   strategy 2

Robust linear program

Summary

$$\min_{x \in \mathbb{R}^n} \ c^\top x \ \text{ s.t. } \ a^\top x \leq b, \ \forall [a^\top \ b] \in \left\{ [a_0^\top \ b_0] + \sum_{l=1}^{k} \zeta_l \, [a_l^\top \ b_l] : \zeta \in Z \subseteq \mathbb{R}^k \right\}$$

Uncertainty Set Z	Convex reformulation
Linear	LP
SOC	SOCP
Conic	Conic program

Robust second-order cone program

Consider

$$\min_{x \in \mathbb{R}^n} \quad c^{\mathsf{T}} x \quad \text{s.t.} \quad \|A(\zeta)x + b(\zeta)\|_2 \ \le\ \alpha^{\mathsf{T}}(\zeta)x + \beta(\zeta), \qquad \forall \zeta \in Z \subseteq \mathbb{R}^k$$

where $\big(A(\zeta), b(\zeta)\big)$ and $\big(\alpha(\zeta), \beta(\zeta)\big)$ are affine functions of ζ:

$$A(\zeta) := A_0 + \sum_{l=1}^{k} \zeta_l A_l \ \in\ \mathbb{R}^{m \times n}, \qquad b(\zeta) := b_0 + \sum_{l=1}^{k} \zeta_l b_l \ \in\ \mathbb{R}^m$$

$$\alpha(\zeta) = \alpha_0 + \sum_{l=1}^{k} \zeta_l \alpha_l \ \in\ \mathbb{R}^n, \qquad \beta(\zeta) := \beta_0 + \sum_{l=1}^{k} \zeta_l \beta_l \ \in\ \mathbb{R}$$

$(A_l, b_l, \alpha_l, \beta_l, l \ge 0)$ are fixed and given; ζ is the uncertain parameter

Formulation is general and allows each entry of the nominal $(A_0, b_0, \alpha_0, \beta_0)$ to be perturbed independently

Robust second-order cone program

Consider

$$\min_{x \in \mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad \|A(\zeta)x + b(\zeta)\|_2 \le \alpha^\mathsf{T}(\zeta)x + \beta(\zeta), \qquad \forall \zeta \in Z \subseteq \mathbb{R}^k$$

where $\big(A(\zeta), b(\zeta)\big)$ and $\big(\alpha(\zeta), \beta(\zeta)\big)$ are affine functions of ζ:

$$A(\zeta) := A_0 + \sum_{l=1}^{k} \zeta_l A_l \in \mathbb{R}^{m \times n}, \qquad b(\zeta) := b_0 + \sum_{l=1}^{k} \zeta_l b_l \in \mathbb{R}^m$$

$$\alpha(\zeta) = \alpha_0 + \sum_{l=1}^{k} \zeta_l \alpha_l \in \mathbb{R}^n, \qquad \beta(\zeta) := \beta_0 + \sum_{l=1}^{k} \zeta_l \beta_l \in \mathbb{R}$$

Generally intractable, except e.g. $Z = \text{conv}(\zeta^1, \ldots, \zeta^p) \subseteq \mathbb{R}^k$ in which case the semi-infinite set of constraints reduces to

$$\|A(\zeta^i)x + b(\zeta^i)\|_2 \le \alpha^\mathsf{T}(\zeta^i)x + \beta(\zeta^i), \quad i = 1, \ldots, p$$

Robust second-order cone program

Decoupled constraints

Special case: left-hand side uncertainty ζ^{l} and right-hand side uncertainty ζ^{r} are decoupled:

$$\min_{x \in \mathbb{R}^n} \quad c^{\mathsf{T}} x \quad \text{s.t.} \quad \left\| A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}}) \right\|_2 \leq \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}}), \quad \forall \zeta^{\mathsf{l}} \in Z^{\mathsf{l}},\ \zeta^{\mathsf{r}} \in Z^{\mathsf{r}}$$

$x \in \mathbb{R}^n$ if feasible iff $\exists \tau$ s.t.

$$\max_{\zeta^{\mathsf{l}} \in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \quad \leq \quad \tau \quad \leq \quad \min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})\, x + \beta(\zeta^{\mathsf{r}})$$

Two classes of uncertainty sets $(Z^{\mathsf{l}}, Z^{\mathsf{r}})$ for which both maximization and minimization have finite convex representations, and hence robust SOCP is tractable

Robust second-order cone program

Interval + conic uncertainties

1. Left-side uncertainty: $A(\zeta^{\mathrm{l}}) = A_0 + \Delta A$ and $b(\zeta^{\mathrm{l}}) = b_0 + \Delta b$ with

$$Z^{\mathrm{l}} := \left\{ \zeta^{\mathrm{l}} := [\Delta A \;\; \Delta b] : |\Delta A_{ij}| \leq \delta_{ij},\ |\Delta b_i| \leq \delta_i,\ i = 1,\ldots,m,\ j = 1,\ldots,n \right\}$$

Subproblem: $\max_{\zeta^{\mathrm{l}} \in Z^{\mathrm{l}}} \|A(\zeta^{\mathrm{l}})x + b(\zeta^{\mathrm{l}})\|_2 \;\leq\; \tau$ (strategy 1: solve in closed form)

2. Right-side uncertainty: $\alpha(\zeta^{\mathrm{r}}) := \alpha_0 + \sum_{l=1}^{k_{\mathrm{r}}} \zeta_l \alpha_l \in \mathbb{R}^n$ and $\beta(\zeta^{\mathrm{r}}) := \beta_0 + \sum_{l=1}^{k_{\mathrm{r}}} \zeta_l \beta_l \in \mathbb{R}$ with

$$Z^{\mathrm{r}} := \left\{ \zeta^{\mathrm{r}} \in \mathbb{R}^{k_{\mathrm{r}}} : \exists u \text{ s.t. } P\zeta^{\mathrm{r}} + Qu + d \in K \right\}$$

Subproblem: $\tau \;\leq\; \min_{\zeta^{\mathrm{r}} \in Z^{\mathrm{r}}} \alpha^{\mathsf{T}}(\zeta^{\mathrm{r}})x + \beta(\zeta^{\mathrm{r}})$ (same as robust LP ⇔ conic constraint)

Suppose Z^{r} satisfies Slater condition: Z^{r} is nonempty and either K is polyhedral or $\exists(\bar{\zeta}^{\mathrm{r}}, \bar{u})$ s.t. $P\bar{\zeta}^{\mathrm{r}} + Q\bar{u} + d \in \mathrm{ri}(K)$

Robust second-order cone program

Interval + conic uncertainties

Theorem

Suppose Z^{r} is nonempty and

- Slater condition: either K is polyhedral cone or $\exists(\bar{\zeta}^{\mathsf{r}}, \bar{u})$ s.t. $P\bar{\zeta}^{\mathsf{r}} + Q\bar{u} + d \in \mathrm{ri}(K)$
- For each x, $\min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \; \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}})$ is finite

Then robust SOCP is equivalent to conic program: $\min_{(x,y,z)} \; c^{\mathsf{T}}x$ s.t.

$$z_i = \left| \sum_j [A_0]_{ij}x_j + [b_0]_i \right| + \sum_j \delta_{ij}|x_j| + \delta_i, \quad i = 1,\ldots,m-1$$

$\Leftrightarrow \max_{\zeta^{\mathsf{l}} \in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \le \tau$

$$\|z\|_2 \le \hat{\beta}(x) - y^{\mathsf{T}}d, \quad y \in K^*, \quad P^{\mathsf{T}}y = \hat{\alpha}(x), \quad Q^{\mathsf{T}}y = 0$$

strong duality — dual feasibility — stationarity

$\Leftrightarrow \tau \le \min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}})$

(same as robust LP)

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

1. Left-side uncertainty: $A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}}) = \left(A_0 x + b_0\right) + L^{\mathsf{T}}(x)\zeta^{\mathsf{l}} r(x)$ with

$$Z^{\mathsf{l}} := \left\{ \zeta^{\mathsf{l}} \in \mathbb{R}^{k_1 \times k_2} : \left\| \zeta^{\mathsf{l}} \right\|_2 := \max_{u : \|u\|_2 \le 1} \left\| \zeta^{\mathsf{l}} u \right\|_2 \le 1 \right\}$$

   At most one of $L(x)$ and $r(x)$ depends on x; moreover dependence is affine in x

   Subproblem: $\max_{\zeta^{\mathsf{l}} \in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \le \tau$ (reduce to LMIs using S-lemma)

2. Right-side uncertainty: same

   Subproblem: $\tau \le \min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}})$ (same as robust LP ⇔ conic constraint)

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

Theorem

Suppose Z^{r} is nonempty and

- Slater condition: either K is polyhedral cone or $\exists(\bar{\zeta}^{\mathrm{r}}, \bar{u})$ s.t. $P\bar{\zeta}^{\mathrm{r}} + Q\bar{u} + d \in \mathrm{ri}(K)$
- For each x, $\min_{\zeta^{\mathrm{r}} \in Z^{\mathrm{r}}} \ \alpha^{\mathsf{T}}(\zeta^{\mathrm{r}})x + \beta(\zeta^{\mathrm{r}})$ is finite

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

Theorem

Then robust SOCP is equivalent to conic program: $\min_{(x,y,\tau,\lambda)} \quad c^\top x$ s.t.

$$y \in K^*, \quad \tau \le \hat{\beta}(x) - y^\top d, \quad P^\top y = \hat{\alpha}(x), \quad Q^\top y = 0$$

$\Leftrightarrow \tau \le \min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \alpha^\top(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}})$

1. If $A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}}) = \left(A_0 x + b_0\right) + L^\top(x)\zeta^{\mathsf{l}} r$ then

$$\lambda \ge 0, \qquad \begin{bmatrix} \tau - \lambda \|r\|_2^2 & (A_0 x + b_0)^\top & 0 \\ A_0 x + b_0 & \tau \mathbb{I}_m & L^\top(x) \\ 0 & L(x) & \lambda \mathbb{I}_{k_1} \end{bmatrix} \succeq 0$$

$\Leftrightarrow \max_{\zeta^{\mathsf{l}} \in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \le \tau$

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

Theorem

Then robust SOCP is equivalent to conic program: $\min\limits_{(x,y,\tau,\lambda)} \quad c^\top x$ s.t.

$$y \in K^*, \quad \tau \le \hat{\beta}(x) - y^\top d, \quad P^\top y = \hat{\alpha}(x), \quad Q^\top y = 0$$

$\Leftrightarrow \tau \le \min\limits_{\zeta^r \in Z^r} \alpha^\top(\zeta^r)x + \beta(\zeta^r)$

2. If $A(\zeta^l)x + b(\zeta^l) = \left(A_0 x + b_0\right) + L^\top \zeta^l r(x)$ then

$$\lambda \ge 0, \qquad \begin{bmatrix} \tau & (A_0 x + b_0)^\top & r^\top(x) \\ A_0 x + b_0 & \tau \mathbb{I}_m - \lambda L^\top L & 0 \\ r(x) & 0 & \lambda \mathbb{I}_{k_2} \end{bmatrix} \succeq 0$$

$\Leftrightarrow \max\limits_{\zeta^l \in Z^l} \|A(\zeta^l)x + b(\zeta^l)\|_2 \le \tau$

Robust semidefinite program

1. Nominal SDP

$$\min_{x \in \mathbb{R}^n} \quad f(x) \quad \text{s.t.} \quad h_0(x) := B_0 + \sum_{i=1}^{n} x_i A_0^i \ \in \ K_{\mathsf{psd}} \ \subset \ \mathbb{S}^m$$

2. Robust SDP

$$\min_{x \in \mathbb{R}^n} \quad f(x) \qquad \text{s.t.} \qquad h_0(x) + h(x, \zeta) \ \in \ K_{\mathsf{psd}}, \qquad \forall \zeta \in Z$$

where

$$h(x, \zeta) \ := \ L^{\mathsf{T}}(x)\zeta R(x) + R^{\mathsf{T}}(x)\zeta^{\mathsf{T}} L(x) \quad \in \quad \mathbb{S}^m$$

$$Z := \left\{ \ \zeta \in \mathbb{R}^{k_1 \times k_2} \ : \ \|\zeta\|_2 := \max_{u : \|u\|_2 = 1} \|\zeta u\|_2 \ \leq \ \rho \ \right\}$$

At most one of the matrices $L(x)$ and $R(x)$ depends on x; moreover dependence is affine in x

to obtain **linear** matrix inequalities

Example: SDP relaxation of OPF

SDP relaxation of OPF:

$$\min_{W \in K_{\text{psd}}} \quad \text{tr}\left(C_0 W\right) \quad \text{s.t.} \quad \text{tr}\left(\Phi_j W\right) \leq p_j^{\max}, \qquad -\text{tr}\left(\Phi_j W\right) \leq -p_j^{\min}$$

$$\text{tr}\left(\Psi_j W\right) \leq q_j^{\max}, \qquad -\text{tr}\left(\Psi_j W\right) \leq -q_j^{\min}$$

$$\text{tr}\left(J_j W\right) \leq v_j^{\max}, \qquad -\text{tr}\left(J_j W\right) \leq -v_j^{\min}$$

where

$$\Phi_j := \frac{1}{2}\left(Y_0^{\mathsf{H}} e_j e_j^{\mathsf{T}} + e_j e_j^{\mathsf{T}} Y_0\right), \qquad \Psi_j := \frac{1}{2i}\left(Y_0^{\mathsf{H}} e_j e_j^{\mathsf{T}} - e_j e_j^{\mathsf{T}} Y_0\right), \qquad J_j := e_j e_j^{\mathsf{T}}$$

and $Y_0 \in \mathbb{C}^{(N+1)\times(N+1)}$ is a given nominal admittance matrix

Φ_j and Ψ_j depend on admittance matrix

Example: SDP relaxation of OPF

Nominal SDP: dual problem

$$-\min_{x\in\mathbb{R}^n} \; c^\top x \quad \text{s.t.} \quad x \geq 0,\; h_0(x) \in K_{\mathsf{psd}}$$

where

$$h_0(x) := \; C_0 + \sum_{i=1}^{N+1} \Bigg(\left(x_{2i-1} - x_{2i}\right) \Phi_i \; + \; \left(x_{2(N+1)+2i-1} - x_{2(N+1)+2i}\right) \Psi_i$$

$$+ \; \left(x_{4(N+1)+2i-1} - x_{4(N+1)+2i}\right) J_i \Bigg)$$

which is in standard form: $\min_{x\in\mathbb{R}^n} \; f(x) \quad \text{s.t.} \quad h_0(x) := B_0 + \sum_{i=1}^{n} x_i A_0^i \; \in \; K_{\mathsf{psd}}$

Example: SDP relaxation of OPF

Uncertain admittance matrix $Y = Y_0 + \Delta Y$

Uncertainty: admittance matrix $Y = Y_0 + \Delta Y$

This results in uncertainty in $h(x)$:

$$h(x, \Delta Y) := L^{\mathsf{H}}(x)\Delta Y + \Delta Y^{\mathsf{H}} L(x)$$

$$L(x) := \sum_{i=1}^{N+1} \left(\frac{1}{2}\left(x_{2i-1} - x_{2i}\right) + \frac{1}{2i}\left(x_{2(N+1)+2i-1} - x_{2(N+1)+2i}\right)\right) e_i e_i^{\mathsf{T}}$$

Robust SDP:

$$-\min_{x \in \mathbb{R}^n} c^{\mathsf{T}} x \quad \text{s.t.} \quad x \geq 0,\ h_0(x) + h(x, \Delta Y) \in K_{\text{psd}}$$

which is in standard form with $h_\zeta(x) := L^{\mathsf{T}}(x)\zeta\mathbb{1} + \mathbb{1}\zeta^{\mathsf{T}} L(x)$

Robust semidefinite program

Theorem

Robust SDP is equivalent to SDP: $\min_{(x,\lambda)} \quad f(x)$ s.t.

1. If $h_\zeta(x) := L^\top(x)\zeta R + R^\top \zeta^\top L(x)$ then

$$\lambda \geq 0, \qquad \begin{bmatrix} h_0(x) - \lambda R^\top R & \rho L^\top(x) \\ \rho L(x) & \lambda \mathbb{I}_{k_1} \end{bmatrix} \succeq 0$$

2. $h_\zeta(x) := L^\top \zeta R(x) + R^\top(x) \zeta^\top L$ then

$$\lambda \geq 0, \qquad \begin{bmatrix} h_0(x) - \lambda L^\top L & \rho R^\top(x) \\ \rho R(x) & \lambda \mathbb{I}_{k_2} \end{bmatrix} \succeq 0$$

Outline

1. Robust optimization
    - General formulation
    - Robust linear program
    - Robust second-order cone program
    - Robust semidefinite program
    - Proofs
2. Chance constrained optimization
3. Convex scenario optimization
4. Stochastic optimization with recourse
5. Applications

Proofs

The proofs illustrate two useful techniques in this, and many other, types of problems

1. Robust LP: conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$
   - Replace subproblem $\bar{h}(x) \leq 0$ by strong duality and KKT condition

strategy 2

2. Robust SOCP: bounded l_2-norm + conic uncertainty
   - Express K_{soc} as K_{psd}
   - Use S-lemma to reduce $\max_{\zeta^l \in Z^l} \|A(\zeta^l)x + b(\zeta^l)\|_2 \leq \tau$ to LMIs

strategy 3

3. S-lemma
   - Use separating hyperplane theorem (similar to Slater theorem proof)

Robust linear program

Consider

$$\min_{x\in\mathbb{R}^n} \; c^\top x \;\; \text{s.t.} \;\; a^\top x \le b, \; \forall [a^\top \; b] \in \left\{ \left[a_0^\top \; b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top \; b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

<u>**Theorem**</u>

3. Conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$ where $K \subseteq \mathbb{R}^m$ is a closed convex pointed cone with nonempty interior. Suppose Z is nonempty and
   - Slater condition: either K is polyhedral cone or $\exists(\bar{\zeta}, \bar{u}) \in \mathbb{R}^{k+p}$ s.t. $P\bar{\zeta} + Q\bar{u} + d \in \text{ri}(K)$
   - For each x, $\max_{\zeta\in Z} \sum_l \zeta_l \left(a_l^\top x - b_l\right)$ is finite

Then (1) is equivalent to conic program: $\min_{(x,y)\in\mathbb{R}^{n+m}} \; c^\top x$ s.t.

strategy 2

$$a_0^\top x + d^\top y \le b_0, \quad y \in K^*, \quad Q^\top y = 0, \; a_l^\top x + \left(P^\top y\right)_l = b_l, \; l = 1,\ldots,k$$

strong duality — dual feasibility — stationarity

Robust linear program

Proof

Recall the subproblem and feasibility condition is:

$$\bar{h}(x) := \max_{\zeta \in Z} \sum_{l=1}^{k} \zeta_l (a_l^\top x - b_l) \ \leq -(a_0^\top x - b_0)$$

Define $s \in \mathbb{R}^k$ by $s_l := s_l(x) := a_l^\top x - b_l$

Then subproblem is:

$$p^*(x) \quad := \quad \max_{(\zeta,u)\in\mathbb{R}^{k+p}} \quad s^\top(x)\zeta \quad \text{s.t.} \quad \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} \zeta \\ u \end{bmatrix} + d \in K$$

Hence the constraint $\bar{h}(x) \leq -(a_0^\top x - b_0)$ is: $p^*(x) \leq -(a_0^\top x - b_0)$

Lagrangian is: for all $(\zeta, u) \in \mathbb{R}^{k+p}$, $y \in K^*$,

$$L(\zeta, u, y) := \ s^\top \zeta + y^\top \left(\begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} \zeta \\ u \end{bmatrix} + d \right) \ = \ y^\top d + \left(s^\top + y^\top P \right) \zeta + y^\top Q u$$

Robust linear program

Proof

stationarity $\nabla_{\zeta,u} L(\zeta, u, y) = 0$

Dual function is:

$$d(y) \;:=\; \max_{(\zeta,u)\in\mathbb{R}^{k+p}} L(\zeta, u, y) \;=\; \begin{cases} d^\top y & \text{if } P^\top y = -s,\ Q^\top y = 0 \\ \infty & \text{otherwise} \end{cases}$$

Dual problem is:

$$d^*(x) \;:=\; \min_{y\in K^*} \; d^\top y \quad \text{s.t.} \quad P^\top y = -s(x),\ Q^\top y = 0$$

Slater Theorem applies (finite optimal primal value, convexity, Slater condition) to conclude strong duality and existence of dual optimal solution $y := y(x)$:

$$p^*(x) = d^*(x) = d^\top y$$

Therefore feasibility $p^*(x) \le -(a_0^\top x - b_0)$ is equivalent to: $d^\top y \le -(a_0^\top x - b_0)$

Robust linear program

Consider

$$\min_{x\in\mathbb{R}^n} \ c^\top x \ \text{ s.t. } \ a^\top x \le b, \ \forall [a^\top \ b] \in \left\{ \left[a_0^\top \ b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top \ b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

Theorem

3. Conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$ where $K \subseteq \mathbb{R}^m$ is a closed convex pointed cone with nonempty interior. Suppose Z is nonempty and
   - Slater condition: either K is polyhedral cone or $\exists(\bar{\zeta}, \bar{u}) \in \mathbb{R}^{k+p}$ s.t. $P\bar{\zeta} + Q\bar{u} + d \in \text{ri}(K)$
   - For each x, $\max_{\zeta\in Z} \sum_l \zeta_l \left(a_l^\top x - b_l\right)$ is finite

   Then (1) is equivalent to conic program: $\min_{(x,y)\in\mathbb{R}^{n+m}} \ c^\top x$ s.t.

$$a_0^\top x + d^\top y \le b_0, \quad y \in K^*, \quad Q^\top y = 0, \ a_l^\top x + \left(P^\top y\right)_l = b_l, \ l = 1,\ldots,k$$

strong duality

Robust linear program

Proof

To ensure $y := y(x)$ is dual optimal, it is necessary and sufficient it satisfies KKT condition for

$$\min_{y \in K^*} \; d^\top y \quad \text{s.t.} \quad P^\top y = -s(x),\; Q^\top y = 0$$

$$\max_{(\zeta, u) \in \mathbb{R}^{k+p}} \; s^\top(x)\zeta \quad \text{s.t.} \quad \begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} \zeta \\ u \end{bmatrix} + d \in K$$

Dual feasibility: $y \in K^*$

Stationarity: $P^\top y = -s(x),\; Q^\top y = 0$

Complementary slackness: $y^\top \left(\begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} \zeta \\ u \end{bmatrix} + d \right) = 0$

this involves ζ : eliminate it using stationarity and strong duality

Robust linear program

Proof

Complementary slackness is implied by stationarity and strong duality:

$$
\begin{aligned}
y^\top \left(\begin{bmatrix} P & Q \end{bmatrix} \begin{bmatrix} \zeta \\ u \end{bmatrix} + d \right) &= \; y^\top P\zeta + y^\top Qu + y^\top d \\
&= -s^\top \zeta + 0 + y^\top d \qquad \text{stationarity: } P^\top y = -s(x),\; Q^\top y = 0 \\
&= \; 0 \qquad \text{strong duality: } s^\top \zeta = d^\top y
\end{aligned}
$$

Robust linear program

Consider

$$\min_{x \in \mathbb{R}^n} \; c^\top x \;\; \text{s.t.} \;\; a^\top x \le b, \; \forall [a^\top \; b] \in \left\{ \left[a_0^\top \; b_0\right] + \sum_{l=1}^{k} \zeta_l \left[a_l^\top \; b_l\right] : \zeta \in Z \subseteq \mathbb{R}^k \right\} \tag{1}$$

Theorem

3. Conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$ where $K \subseteq \mathbb{R}^m$ is a closed convex pointed cone with nonempty interior. Suppose Z is nonempty and
   - Slater condition: either K is polyhedral cone or $\exists (\bar{\zeta}, \bar{u}) \in \mathbb{R}^{k+p}$ s.t. $P\bar{\zeta} + Q\bar{u} + d \in \text{ri}(K)$
   - For each x, $\max_{\zeta \in Z} \sum_l \zeta_l \left(a_l^\top x - b_l\right)$ is finite

Then (1) is equivalent to conic program: $\min_{(x,y) \in \mathbb{R}^{n+m}} \; c^\top x$ s.t.

$$a_0^\top x + d^\top y \le b_0, \;\; y \in K^*, \;\; Q^\top y = 0, \; a_l^\top x + \left(P^\top y\right)_l = b_l, \; l = 1, \ldots, k$$

dual feasibility | stationarity

Proofs

1. Robust LP: conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p \text{ s.t. } P\zeta + Qu + d \in K\}$
    - Replace subproblem $\bar{h}(x) \leq 0$ by strong duality and KKT condition

2. Robust SOCP: bounded l_2-norm + conic uncertainty
    - Express K_{soc} as K_{psd}
    - Use S-lemma to reduce $\max_{\zeta^{\text{l}} \in Z^{\text{l}}} \|A(\zeta^{\text{l}})x + b(\zeta^{\text{l}})\|_2 \leq \tau$ as LMIs

strategy 3

3. S-lemma
    - Use separating hyperplane theorem (similar to Slater theorem proof)

Robust second-order cone program

Decoupled constraints

Special case: left-hand side uncertainty ζ^{l} and right-hand side uncertainty ζ^r are decoupled:

$$\min_{x\in\mathbb{R}^n} \quad c^{\mathsf{T}}x \quad \text{s.t.} \quad \left\| A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}}) \right\|_2 \leq \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}}), \quad \forall \zeta^{\mathsf{l}} \in Z^{\mathsf{l}},\ \zeta^{\mathsf{r}} \in Z^{\mathsf{r}}$$

$x \in \mathbb{R}^n$ if feasible iff $\exists \tau$ s.t.

$$\max_{\zeta^{\mathsf{l}}\in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \quad \leq \quad \tau \quad \leq \quad \min_{\zeta^{\mathsf{r}}\in Z^{\mathsf{r}}} \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})\, x + \beta(\zeta^{\mathsf{r}})$$

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

1. Left-side uncertainty: $A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}}) = \left(A_0 x + b_0\right) + L^{\mathsf{T}}(x)\zeta^{\mathsf{l}} r(x)$ with

$$Z^{\mathsf{l}} := \left\{ \zeta^{\mathsf{l}} \in \mathbb{R}^{k_1 \times k_2} : \left\| \zeta^{\mathsf{l}} \right\|_2 := \max_{u : \|u\|_2 \le 1} \left\| \zeta^{\mathsf{l}} u \right\|_2 \le 1 \right\}$$

At most one of $L(x)$ and $r(x)$ depends on x; moreover dependence is affine in x

Subproblem: $\max_{\zeta^{\mathsf{l}} \in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \le \tau$ (reduce to LMIs using S-lemma)

2. Right-side uncertainty: same

Subproblem: $\tau \le \min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}})$ (same as robust LP ⇔ conic constraint)

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

Theorem

Suppose Z^{r} is nonempty and

- Slater condition: either K is polyhedral cone or $\exists(\bar{\zeta}^{\mathsf{r}}, \bar{u})$ s.t. $P\bar{\zeta}^{\mathsf{r}} + Q\bar{u} + d \in \mathrm{ri}(K)$
- For each x, $\min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \ \alpha^{\mathsf{T}}(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}})$ is finite

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

Theorem

Then robust SOCP is equivalent to conic program: $\min_{(x,y,\tau,\lambda)} \quad c^\top x$ s.t.

$$y \in K^*, \quad \tau \le \hat{\beta}(x) - y^\top d, \quad P^\top y = \hat{\alpha}(x), \quad Q^\top y = 0$$

$\Leftrightarrow \tau \le \min_{\zeta^r \in Z^r} \alpha^\top(\zeta^r)x + \beta(\zeta^r)$

1. If $A(\zeta^l)x + b(\zeta^l) = \left(A_0 x + b_0\right) + L^\top(x)\zeta^l r$ then

$$\lambda \ \geq \ 0, \qquad \begin{bmatrix} \tau - \lambda\|r\|_2^2 & (A_0 x + b_0)^\top & 0 \\ A_0 x + b_0 & \tau \mathbb{I}_m & L^\top(x) \\ 0 & L(x) & \lambda \mathbb{I}_{k_1} \end{bmatrix} \succeq 0$$

$\Leftrightarrow \max_{\zeta^l \in Z^l} \|A(\zeta^l)x + b(\zeta^l)\|_2 \le \tau$

Robust second-order cone program

Bounded ℓ_2 norm + conic uncertainties

Theorem

Then robust SOCP is equivalent to conic program: $\min_{(x,y,\tau,\lambda)} \quad c^\top x$ s.t.

$$y \in K^*, \quad \tau \leq \hat{\beta}(x) - y^\top d, \quad P^\top y = \hat{\alpha}(x), \quad Q^\top y = 0$$

$\Leftrightarrow \tau \leq \min_{\zeta^{\mathsf{r}} \in Z^{\mathsf{r}}} \alpha^\top(\zeta^{\mathsf{r}})x + \beta(\zeta^{\mathsf{r}})$

2. If $A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}}) = \left(A_0 x + b_0\right) + L^\top \zeta^{\mathsf{l}} r(x)$ then

$$\lambda \geq 0, \qquad \begin{bmatrix} \tau & (A_0 x + b_0)^\top & r^\top(x) \\ A_0 x + b_0 & \tau I_m - \lambda L^\top L & 0 \\ r(x) & 0 & \lambda I_{k_2} \end{bmatrix} \succeq 0$$

$\Leftrightarrow \max_{\zeta^{\mathsf{l}} \in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \leq \tau$

Robust second-order cone program

Proof

Prove $\max_{\zeta^{\mathsf{l}} \in Z^{\mathsf{l}}} \|A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}})\|_2 \leq \tau$ is equivalent to LMIs:

1. If $A(\zeta^{\mathsf{l}})x + b(\zeta^{\mathsf{l}}) = \left(A_0 x + b_0\right) + \; L^{\mathsf{T}}(x)\zeta^{\mathsf{l}} r$ then

$$\lambda \;\geq\; 0, \qquad \begin{bmatrix} \tau - \lambda\|r\|_2^2 & (A_0 x + b_0)^{\mathsf{T}} & 0 \\ A_0 x + b_0 & \tau \mathbb{I}_m & L^{\mathsf{T}}(x) \\ 0 & L(x) & \lambda \mathbb{I}_{k_1} \end{bmatrix} \;\succeq\; 0$$

Robust second-order cone program

Proof

3 ideas:

1. K_{soc} as K_{psd} : $(y, t) \in K_{\text{soc}}$, i.e., $\|y\|_2 \leq t$ if and only if $\begin{bmatrix} t & y^\top \\ y & t\mathbb{I}_l \end{bmatrix} \succeq 0$

2. l_2-norm matrix minimization : $-\rho\|a_1\|_2\|a_2\|_2 \;=\; \min_{X:\|X\|_2\leq\rho} a_1^\top X a_2$

3. S-lemma : Suppose $\bar{x}^\top A\bar{x} > 0$ for some $\bar{x}$. Then $x^\top Ax \geq 0 \Rightarrow x^\top Bx \geq 0$ holds if and only if $B \succeq \lambda A$ for some $\lambda \geq 0$

Robust second-order cone program

Proof

Let $g(x) := A_0 x + b_0 \in \mathbb{R}^m$

Subproblem $\max_{\zeta^l \in Z^l} \|A(\zeta^l)x + b(\zeta^l)\|_2 \le \tau$ is equivalent to:

$$\begin{bmatrix} \tau & \left(g(x) + L^\top(x)\zeta^l r\right)^\top \\ g(x) + L^\top(x)\zeta^l r & \tau \mathbb{I}_m \end{bmatrix} \succeq 0, \qquad \zeta^l \in Z^l$$

Or:

$$(z_1)^2\tau \;+\; 2z_2^\top \left(g(x) + L^\top(x)\zeta^l r\right) z_1 \;+\; (z_2^\top z_2)\tau \;\ge\; 0, \quad \forall z_1 \in \mathbb{R},\, z_2 \in \mathbb{R}^m,\, \zeta^l \in Z^l$$

Or:

$$(z_1)^2\tau + 2z_2^\top g(x) z_1 + (z_2^\top z_2)\tau + \min_{\zeta^l : \|\zeta^l\|_2 \le 1} (2L(x)z_2)^\top \zeta^l (z_1 r) \;\ge\; 0 \quad \forall z_1 \in \mathbb{R},\, z_2 \in \mathbb{R}^m$$

Robust second-order cone program

Proof

Apply l_2-norm matrix minimization twice:

$$\min_{\zeta^{\mathsf{I}}:\|\zeta^{\mathsf{I}}\|_2\le 1} (2L(x)z_2)^\mathsf{T}\zeta^{\mathsf{I}}(z_1 r) = -2\|L(x)z_2\|_2\|z_1 r\|_2 = \min_{X:\|X\|_2\le\|z_1 r\|_2} (2L(x)z_2)^\mathsf{T}X(1)$$

Therefore, for all $z_1 \in \mathbb{R}\$, z_2 \in \mathbb{R}^m, X \in \mathbb{R}^{k_1}$, if $z_1^2\|r\|_2^2 - X^\mathsf{T}X \ge 0$ then

$$(z_1)^2\tau + 2z_2^\mathsf{T}g(x)z_1 + (z_2^\mathsf{T}z_2)\tau + 2X^\mathsf{T}L(x)z_2 \ge 0$$

This is equivalent to:

$$\begin{bmatrix} \|r\|_2^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\mathbb{I}_{k_1} \end{bmatrix} \succeq 0 \quad \Longrightarrow \quad \begin{bmatrix} \tau & g^\mathsf{T}(x) & 0 \\ g(x) & \tau\mathbb{I}_m & L^\mathsf{T}(x) \\ 0 & L(x) & 0 \end{bmatrix} \succeq 0$$

Robust second-order cone program

Proof

Clearly there exists $z_1 > 0$ such that $z_1^2 \|r\|_2^2 > 0$

Hence S-lemma implies: $\exists \lambda \geq 0$ such that

$$\begin{bmatrix} \tau - \lambda \|r\|_2^2 & g^\mathsf{T}(x) & 0 \\ g(x) & \tau \mathbb{I}_m & L^\mathsf{T}(x) \\ 0 & L(x) & \lambda \mathbb{I}_{k_1} \end{bmatrix} \succeq 0$$

Proofs

1. Robust LP: conic uncertainty $Z := \{\zeta \in \mathbb{R}^k : \exists u \in \mathbb{R}^p$ s.t. $P\zeta + Qu + d \in K\}$
   - Replace subproblem $\bar{h}(x) \leq 0$ by strong duality and KKT condition

2. Robust SOCP: bounded l_2-norm + conic uncertainty
   - Express K_{soc} as K_{psd}
   - Use S-lemma to reduce $\max_{\zeta^{\text{l}} \in Z^{\text{l}}} ||A(\zeta^{\text{l}})x + b(\zeta^{\text{l}})||_2 \leq \tau$ as LMIs

3. S-lemma
   - Use separating hyperplane theorem (similar to Slater theorem proof)

S-lemma

Proof

S-lemma

Let A, B be $n \times n$ symmetric matrices and $\bar{x}^\top A \bar{x} > 0$ for some $\bar{x} \in \mathbb{R}^n$

The following are equivalent

(i) $x^\top A x \geq 0 \Rightarrow x^\top B x \geq 0$

(ii) $\exists \lambda \geq 0$ such that $B \succeq \lambda A$

S-lemma
Proof

S-lemma

Let A, B be $n \times n$ symmetric matrices and $\bar{x}^\top A \bar{x} > 0$ for some $\bar{x} \in \mathbb{R}^n$

The following are equivalent

(i) $x^\top A x \geq 0 \Rightarrow x^\top B x \geq 0$

(ii) $\exists \lambda \geq 0$ such that $B \succeq \lambda A$

Proof

(ii) $\implies$ (i) : $x^\top B x - x^\top \lambda A x = x^\top (B - \lambda A) x \geq 0$. Hence (ii) $\implies$ (i)

S-lemma

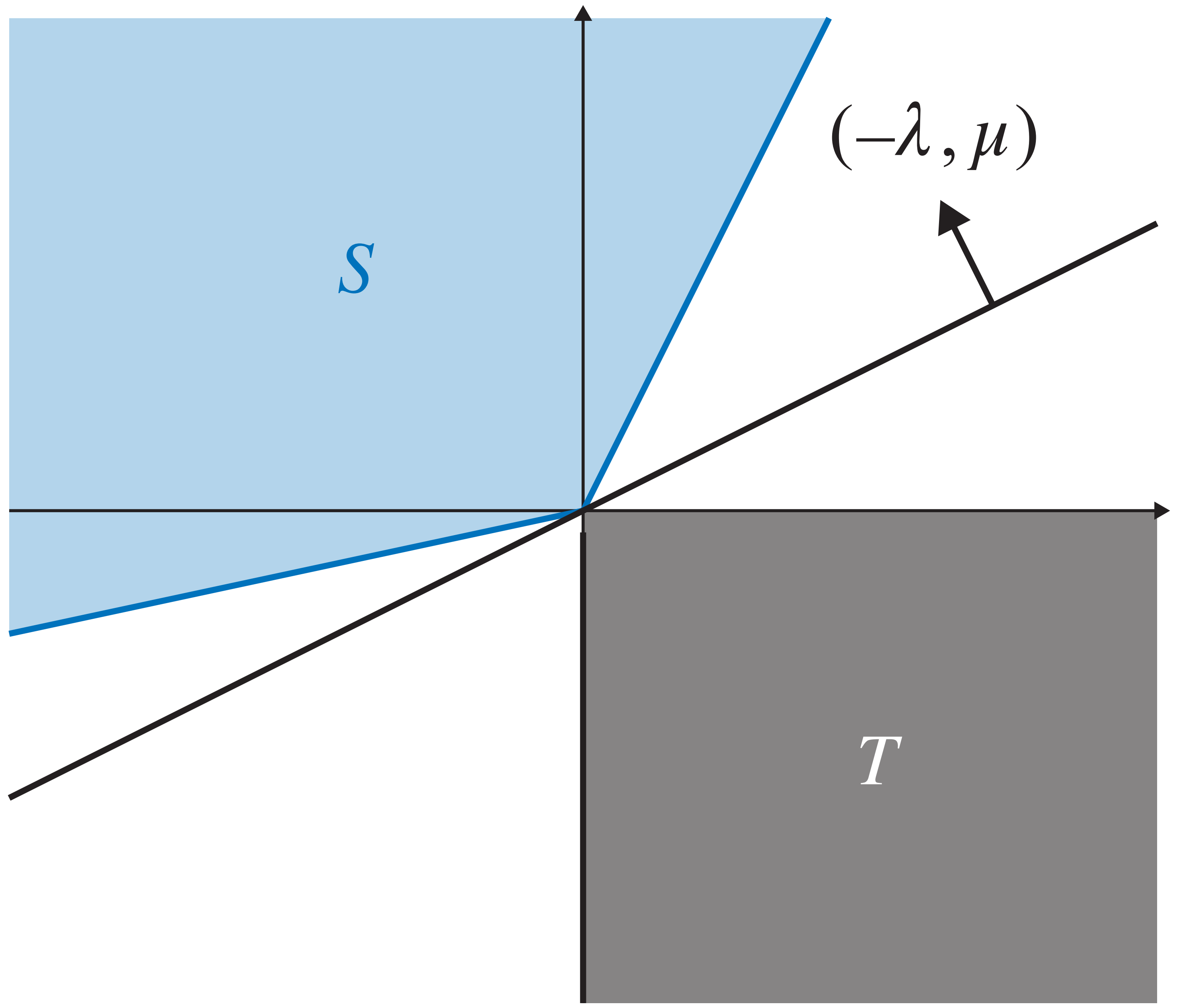
Proof: (i) $\implies$ (ii)

Consider

$$S := \left\{ \begin{bmatrix} x^\top A x \\ x^\top B x \end{bmatrix} \in \mathbb{R}^2 : x \in \mathbb{R}^n \right\}, \qquad T := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 : u \geq 0,\, v < 0 \right\}$$

Will prove in 4 steps:

1. Show that $S \cap T = \emptyset$
2. Show that S is a cone.
3. Show that S is convex.
4. Use the Separating Hyperplane theorem to prove (ii)



The result is shown in the figure

S-lemma

Proof: (i) $\implies$ (ii)

Let

$$\begin{bmatrix} u(x) \\ v(x) \end{bmatrix} := \begin{bmatrix} x^\top Ax \\ x^\top Bx \end{bmatrix} \in S \qquad \text{for all } x \in \mathbb{R}^n$$

Suppose (i) holds.

S-lemma

Proof: (i) $\implies$ (ii)

Let

$$\begin{bmatrix} u(x) \\ v(x) \end{bmatrix} := \begin{bmatrix} x^\top A x \\ x^\top B x \end{bmatrix} \in S \qquad \text{for all } x \in \mathbb{R}^n$$

Suppose (i) holds.

1. $S \cap T = \emptyset$: Since $u(x) \geq 0 \Rightarrow v(x) \geq 0$, we have $(u(x), v(x)) \notin T$. Conversely, if $(a, b) \in T$, then there is no $x \in \mathbb{R}^n$ with $(u(x), v(x)) = (a, b)$

S-lemma

Proof: (i) $\implies$ (ii)

Let

$$\begin{bmatrix} u(x) \\ v(x) \end{bmatrix} := \begin{bmatrix} x^\top A x \\ x^\top B x \end{bmatrix} \in S \qquad \text{for all } x \in \mathbb{R}^n$$

Suppose (i) holds.

1. $S \cap T = \varnothing$: Since $u(x) \geq 0 \Rightarrow v(x) \geq 0$, we have $(u(x), v(x)) \notin T$. Conversely, if $(a, b) \in T$, then there is no $x \in \mathbb{R}^n$ with $(u(x), v(x)) = (a, b)$

2. S is a cone : If $(u(x), v(x)) \in S$, then for any $\lambda^2 > 0$ we have

$$\lambda^2 \begin{bmatrix} u(x) \\ v(x) \end{bmatrix} = \begin{bmatrix} (\lambda x)^\top A (\lambda x) \\ (\lambda x)^\top B (\lambda x) \end{bmatrix} = \begin{bmatrix} u(\lambda x) \\ v(\lambda x) \end{bmatrix} \in S$$

S-lemma

Proof: S is convex

3. S is convex : Let $y_1 := (u(x_1), v(x_1))$ and $y_2 := (u(x_2), v(x_2))$ be in S. Fix any $\alpha \in (0,1)$

Case 1: y_1, y_2 are linearly dependent.

Then $y_1 = cy_2$ for some $c \neq 0$, i.e., y_1, y_2 are are on the same ray from 0

Note that $z \;:=\; \alpha y_1 + (1-\alpha)y_2 \;=\; (c\alpha + (1-\alpha))y_2 \;=\; \left(\dfrac{c\alpha + (1-\alpha)}{c}\right) y_1$

i.e., z is on the same ray as y_1 and y_2, and hence must be in S

S-lemma
Proof: S is convex

Case 2: y_1, y_2 are linearly independent, i.e., they form a basis of $\mathbb{R}^2$.

We have to show: $\exists \bar{x} \in \mathbb{R}^n$ such that

$$\begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} \;=\; \alpha y_1 + (1-\alpha) y_2$$

which implies that $z := \alpha y_1 + (1-\alpha) y_2 \;\in\; S$

S-lemma

Proof: S is convex

Case 2: y_1, y_2 are linearly independent, i.e., they form a basis of $\mathbb{R}^2$.

We have to show: $\exists \bar{x} \in \mathbb{R}^n$ such that

$$\begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} = \alpha y_1 + (1-\alpha) y_2$$

which implies that $z := \alpha y_1 + (1-\alpha) y_2 \in S$

Since S is a cone, it suffices to construct $\bar{x}$ such that

$$\begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} = \lambda(\alpha y_1 + (1-\alpha) y_2), \qquad \text{for some } \lambda > 0$$

We will seek $\bar{x}$ of the form $\bar{x} = \alpha x_1 + \beta x_2$, i.e., derive $\beta \in \mathbb{R}$ such that the above holds

S-lemma

Proof: S is convex

Case 2: y_1, y_2 are linearly independent, i.e., they form a basis of $\mathbb{R}^2$.

By definition of $(u(x), v(x))$:

$$\begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} = \begin{bmatrix} (\alpha x_1 + \beta x_2)^\top A(\alpha x_1 + \beta x_2) \\ (\alpha x_1 + \beta x_2)^\top B(\alpha x_1 + \beta x_2) \end{bmatrix} = \begin{bmatrix} \alpha^2 u(x_1) + \beta^2 u(x_2) + 2\alpha\beta x_1^\top A x_2 \\ \alpha^2 v(x_1) + \beta^2 v(x_2) + 2\alpha\beta x_1^\top B x_2 \end{bmatrix}$$

$$= \alpha^2 y_1 + \beta^2 y_2 + 2\alpha\beta \begin{bmatrix} x_1^\top A x_2 \\ x_1^\top B x_2 \end{bmatrix}$$

uses $A^\top = A$, $B^\top = B$

S-lemma

Proof: S is convex

Case 2: y_1, y_2 are linearly independent, i.e., they form a basis of $\mathbb{R}^2$.

By definition of $(u(x), v(x))$:

$$\begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} = \begin{bmatrix} (\alpha x_1 + \beta x_2)^\top A(\alpha x_1 + \beta x_2) \\ (\alpha x_1 + \beta x_2)^\top B(\alpha x_1 + \beta x_2) \end{bmatrix} = \begin{bmatrix} \alpha^2 u(x_1) + \beta^2 u(x_2) + 2\alpha\beta x_1^\top A x_2 \\ \alpha^2 v(x_1) + \beta^2 v(x_2) + 2\alpha\beta x_1^\top B x_2 \end{bmatrix}$$

$$= \alpha^2 y_1 + \beta^2 y_2 + 2\alpha\beta \begin{bmatrix} x_1^\top A x_2 \\ x_1^\top B x_2 \end{bmatrix}$$

uses $A^\top = A$, $B^\top = B$

Since y_1, y_2 form a basis of $\mathbb{R}^2$, we can express $\begin{bmatrix} x_1^\top A x_2 \\ x_1^\top B x_2 \end{bmatrix} =: a y_1 + b y_2$ for some $a, b \in \mathbb{R}$

$$\implies \begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} = (\alpha + 2a\beta)\left(\alpha y_1 + \frac{\beta^2 + 2\alpha b\beta}{\alpha + 2a\beta} y_2 \right)$$

S-lemma
Proof: S is convex

Case 2: y_1, y_2 are linearly independent, i.e., they form a basis of $\mathbb{R}^2$.

Therefore we seek $\beta \in \mathbb{R}$ such that

$$\begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} = (\alpha + 2a\beta)\left(\alpha y_1 + \frac{\beta^2 + 2\alpha b\beta}{\alpha + 2a\beta} y_2 \right) = \lambda(\alpha y_1 + (1-\alpha) y_2), \quad \text{for some } \lambda > 0$$

i.e. we seek $\beta \in \mathbb{R}$ such that

$$\alpha + 2a\beta > 0, \quad \beta^2 + 2\alpha b\beta = (1-\alpha)(\alpha + 2a\beta)$$

S-lemma
Proof: S is convex

Case 2: y_1, y_2 are linearly independent, i.e., they form a basis of $\mathbb{R}^2$.

Therefore we seek $\beta \in \mathbb{R}$ such that

$$\begin{bmatrix} u(\bar{x}) \\ v(\bar{x}) \end{bmatrix} = (\alpha + 2a\beta)\left(\alpha y_1 + \frac{\beta^2 + 2\alpha b\beta}{\alpha + 2a\beta} y_2\right) = \lambda(\alpha y_1 + (1-\alpha)y_2), \quad \text{for some } \lambda > 0$$

i.e. we seek $\beta \in \mathbb{R}$ such that

$$\alpha + 2a\beta > 0, \quad \beta^2 + 2\alpha b\beta = (1-\alpha)(\alpha + 2a\beta)$$

The quadratic equation has two routes, one > 0 and the other < 0

Choose root β such that $a\beta \geq 0$, so that $\alpha + 2a\beta > 0$

This shows $z := \alpha y_1 + (1-\alpha)y_2 \in S$, i.e., S is convex

S-lemma

Proof: (i) $\implies$ (ii)

4. Since S and T are convex and disjoint, the Separating Hyperplane theorem implies there exists nonzero $(-\lambda, \mu) \in \mathbb{R}^2$ sum that

$$-\lambda u + \mu v \;\; \geq -\lambda a + \mu b, \qquad \forall (u, v) \in S,\ (a, b) \in T$$

- Since $0 \in S$, we have $-\lambda a + \mu b \leq 0$ for all $\forall (a, b) \in T$
- This implies $\lambda \geq 0$ and $\mu \geq 0$
- Taking $(a, b) \to 0$, we have $-\lambda u + \mu v \geq 0$ for all $(u, v) \in S$, i.e.,

$$-\lambda x^\top A x + \mu x^\top B x \geq 0 \;\text{ for all } x \in \mathbb{R}^n$$

- If $\mu = 0$, then $\lambda > 0$ (since $(-\lambda, \mu) \neq 0$), but this contradicts the above at $\bar{x}$
- Hence, can take $\mu = 1$, leading to $x^\top B x \geq \lambda x^\top A x$ for all $x \in \mathbb{R}^n$

Outline

1. Robust optimization
2. Chance constrained optimization
    - Tractable instances
    - Concentration inequalities
3. Convex scenario optimization
4. Stochastic optimization with recourse

Chance constrained optimization

Separable constraints

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad \mathbb{P}\left(\zeta \leq h(x)\right) \geq p$$

- $c : \mathbb{R}^n \to \mathbb{R}$: cost function
- $h_i : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$: constraint functions
- ζ : random vector
- $\mathbb{P}$: probability measure
- $p \in [0,1]$
- $X \subseteq \mathbb{R}^n$: nonempty convex

Less conservative than robust optimization and allows constraint violation with probability $< 1 - p$

Chance constrained optimization

Separable constraints

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad \mathbb{P}\left(\zeta \leq h(x)\right) \geq p$$

where $h : \mathbb{R}^n \to \mathbb{R}^m$, $\zeta \in \mathbb{R}^m$

Can express it terms of distribution function F_ζ:

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad F_\zeta(h(x)) \geq p$$

Chance constrained optimization

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad F_\zeta(h(x)) \geq p$$

Will introduce two techniques to deal with chance constrained opt

1. Tractable instances
   - … when constraint functions h_i and probability measure $\mathbb{P}$ have certain concavity properties
   - Study conditions for feasible set to be convex and for strong duality and dual optimality

2. Safe approximation through concentration inequalities
   - Safe approximation: more conservative but simpler to solve
   - Upper bounding violation probability using concentration inequality (e.g. Chernoff bound)
   - Upper bounding distribution of ζ by known distribution (e.g. sub-Gaussian)

Tractable instances

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad F_\zeta(h(x)) \geq p$$

Two equivalent formulations

1. Hides constraint function h and distribution F_ζ in the feasible set X_p

$$\min_{x \in X} c(x) \quad \text{s.t.} \quad x \in X_p \qquad \text{where} \qquad X_p := \left\{ x \in \mathbb{R}^n : F_\zeta(h(x)) \geq p \right\}$$

   - When is X_p a convex set?

Tractable instances

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad F_\zeta(h(x)) \geq p$$

Two equivalent formulations for convexity analysis

1. Hides constraint function h and distribution F_ζ in the feasible set X_p

$$\min_{x \in X} c(x) \quad \text{s.t.} \quad x \in X_p \qquad \text{where} \qquad X_p := \left\{ x \in \mathbb{R}^n : F_\zeta(h(x)) \geq p \right\}$$

- When is X_p a convex set?

2. Characterizes optimality in terms of h through p-level set Z_p of distribution function F_ζ

$$\min_{(x,z) \in X \times Z_p} \quad c(x) \quad \text{s.t.} \quad h(x) \geq z \qquad \text{where} \qquad Z_p := \left\{ z \in \mathbb{R}^m : F_\zeta(z) \geq p \right\}$$

- What are conditions for strong duality and saddle point optimality?

explicit constraint for opt cond

α-concavity

Definition

Let $\Omega \subseteq \mathbb{R}^m$ be a convex set. A nonnegative function $f : \Omega \to \mathbb{R}_+$ is α-concave with $\alpha \in [-\infty, \infty]$ if for all $x, y \in \Omega$ such that $f(x) > 0, f(y) > 0$ and all $\lambda \in [0,1]$, we have

f^α is concave

$$f(\lambda x + (1-\lambda)y) \geq m_\alpha(f(x), f(y), \lambda) := \begin{cases} \left(\lambda f^\alpha(x) + (1-\lambda) f^\alpha(y)\right)^{1/\alpha} & \text{if } \alpha \notin \{0, -\infty, \infty\} \\ f^\lambda(x) f^{1-\lambda}(y) & \text{if } \alpha = 0 \\ \min\{f(x), f(y)\} & \text{if } \alpha = -\infty \\ \max\{f(x), f(y)\} & \text{if } \alpha = \infty \end{cases}$$

- ∞-concavity: constant function f
- 1-concavity: concave
- 0-concavity: log-concave
- $-\infty$-concavity: quasi-concave

α-concavity

Lemma

Consider a convex set $\Omega \subseteq \mathbb{R}^m$ and a nonnegative function $f : \Omega \to \mathbb{R}_+$.

1. The mapping $\alpha \to m_\alpha(a, b, \lambda)$ is nondecreasing in α
2. α-concavity $\Rightarrow$ β-concavity if $\alpha \geq \beta$ (e.g., concavity ⇒ log-concavity ⇒ quasi-concavity)
3. If f is α concave for some $\alpha > -\infty$, then f is continuous in ri(Ω)
4. If all $h_i : \mathbb{R}^n \to \mathbb{R},\ i = 1,\ldots,m,$ are concave and f is nonnegative, nondecreasing and α-concave for some $\alpha \in [-\infty, \infty]$, then $f \circ h : \mathbb{R}^n \to \mathbb{R}_+$ is α-concave
5. Suppose $f := \mathbb{R}^{n_1+n_2}$ is such that, for all $y \in Y \subseteq \mathbb{R}^{n_2}$, $f(x, y)$ is α-concave in x for some $\alpha \in [-\infty, \infty]$ on a convex set $X \subseteq \mathbb{R}^{n_1}$. Then $g(x) := \inf_{y \in Y} f(x, y)$ is α-concave on X

Convexity of X_p

Theorem

Suppose all components h_i of $h : \mathbb{R}^n \to \mathbb{R}^m$ are concave and the distribution function F_ζ is α-concave for some $\alpha \in [-\infty, \infty]$, then the feasible set

$$X_p \;:=\; \left\{ x \in \mathbb{R}^n : F_\zeta(h(x)) \geq p \right\}$$

is closed and convex

Duality and optimality

1. Let p-level set of distribution function $F_\zeta(z)$ be $(p \in (0,1))$

$$Z_p \ := \ \left\{ z \in \mathbb{R}^m : F_\zeta(z) \geq p \right\}$$

distribution F_ζ of ζ is embedded in p-level set Z_p

2. Chance constrained problem is equivalent to:

$$c^* \ := \ \min_{x \in X,\, z \in Z_p} \ c(x) \quad \text{s.t.} \quad h(x) \geq z$$

3. Lagrangian, dual function and dual problem are:

$$L(x, z, \mu) := c(x) + \mu^\top (z - h(x))$$

$$d(\mu) = \underbrace{\inf_{x \in X} \left(c(x) - \mu^\top h(x) \right)}_{d_X(\mu)} + \underbrace{\inf_{z \in Z_p} \mu^\top z}_{d_Z(\mu)}, \qquad \mu \in \mathbb{R}^m$$

$$d^* := \sup_{\mu \geq 0} d(\mu) = \sup_{\mu \geq 0} d_X(\mu) + d_Z(\mu)$$

Duality and optimality

Chance constrained problem and its dual:

$$c^* \quad := \quad \min_{x \in X,\, z \in Z_p} \quad c(x) \quad \text{s.t.} \quad h(x) \geq z$$

$$d^* \quad := \quad \sup_{\mu \geq 0} d_X(\mu) + d_Z(\mu)$$

where $d_X(\mu) := \inf_{x \in X} \left(c(x) - \mu^\top h(x) \right)$ and $d_Z(\mu) := \inf_{z \in Z_p} \mu^\top z$

$d_X(\mu), d_Z(\mu)$ can be extended real-valued and not differentiable, even if c, h are real-valued and differentiable
They are however always concave and hence subdifferentiable

Duality and optimality

Chance constrained problem and its dual:

$$c^* \quad := \min_{x \in X,\, z \in Z_p} \quad c(x) \quad \text{s.t.} \quad h(x) \geq z$$

$$d^* \quad := \sup_{\mu \geq 0} d_X(\mu) + d_Z(\mu)$$

where $d_X(\mu) := \inf_{x \in X} \left(c(x) - \mu^\top h(x) \right)$ and $d_Z(\mu) := \inf_{z \in Z_p} \mu^\top z$

Definition

$(x, z, \mu) \in X \times Z_p \times \mathbb{R}^m_+ \;\subseteq\; \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ is a saddle point if

$$\sup_{\mu \geq 0} L(x^*, z^*, \mu) \;=\; L(x^*, z^*, \mu^*) \;=\; \inf_{(x,z) \in X \times Z_p} L(x, z, \mu^*)$$

Duality and optimality

Assumptions

1. *Convexity*:
    - c is convex; h is concave
    - X is nonempty convex
    - Distribution function $F_\zeta(z)$ is α-concave for an $\alpha \in [-\infty, \infty]$
2. *Slater condition*: one of the following holds
    - CQ1: There exists $(\bar{x}, \bar{z}) \in X \times Z_p$ such that $h(\bar{x}) > \bar{z}$
    - CQ2: Functions h is affine and there exists $(\bar{x}, \bar{z}) \in$ ri$(X \times Z_p)$ such that $h(\bar{x}) \geq \bar{z}$

Duality and optimality

Theorem

Suppose conditions 1 and 2 hold.

1. *Strong duality and optimality*: If $c^* > -\infty$ then $\exists$ dual optimal $\mu^* \geq 0$ that closes the duality gap, i.e., $c^* = d(\mu^*) = d^*$. Moreover the set of dual optima μ^* is convex and closed (compact under CQ1)

2. *Saddle point characterization*: A point $(x^*, z^*, \mu^*) \in X \times Z_p \times \mathbb{R}^m_+$ is primal-dual optimal and closes the duality gap (i.e., $c^* = c(x^*) = d(\mu^*) = d^*$) if and only if

$$d_X(\mu^*) = \ c(x^*) - \mu^{*\mathsf{T}} h(x^*), \qquad d_Z(\mu^*) = \ \mu^{*\mathsf{T}} z^*, \qquad \mu^{*\mathsf{T}}(z^* - h(x^*)) = \ 0$$

Such a point is a saddle point

Primal optimality and dual differentiability

Let primal optima, given μ, be

$$X(\mu) := \{x \in X : d_X(\mu) = c(x) - \mu^\mathsf{T} h(x)\}, \qquad Z(\mu) := \{z \in Z_p : d_Z(\mu) = \mu^\mathsf{T} z\}$$

Theorem holds whether or not $X(\mu), Z(\mu)$ are empty, i.e., primal optimum does not exist

Suppose X, Z_p are nonempty, convex and compact. Then

1. $X(\mu), Z(\mu)$ are nonempty, convex and compact
2. $d(\mu) = d_X(\mu) + d_Z(\mu)$ is real-valued and concave
3. Subdifferentials are

$$\partial d_X(\mu) \;=\; \text{conv}(-h(x) : x \in X(\mu)), \qquad \partial d_Z(\mu) \;=\; Z(\mu)$$

   Hence $\partial d(\mu) = \text{conv}(-h(x) : x \in X(\mu)) + Z(\mu)$

4. Derivative $\nabla d(\mu) = -h(x^*) + z^*$ exists if $X(\mu)), Z(\mu)$ are singletons

Outline

1. Robust optimization
2. Chance constrained optimization
    - Tractable instances
    - Concentration inequalities
3. Convex scenario optimization
4. Stochastic optimization with recourse
5. Applications

Chance constrained optimization

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad F_{\zeta}(h(x)) \geq p$$

Will introduce two techniques to deal with chance constrained opt

1. Tractable instances
    - … when constraint functions h_i and probability measure $\mathbb{P}$ have certain concavity properties
    - Study conditions for feasible set to be convex and for strong duality and dual optimality

2. Safe approximation through concentration inequalities
    - Safe approximation: more conservative but simpler to solve
    - Upper bounding violation probability using concentration inequality (e.g. Chernoff bound)
    - Upper bounding distribution of ζ by known distribution (e.g. sub-Gaussian)

Safe approximation

Example

Chance constrained linear program:

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(\sum_{l=1}^{k} \left(a_l^\top x - b_l\right) \zeta_l \le -(a_0^\top x - b_0) \right) \ \ge \ 1 - \epsilon$$

The following SOCP is a safe approximation:

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad r\|\hat{A}x - \hat{b}\|_2 \le -(\hat{a}_0^\top x - \hat{b}_0)$$

where $\hat{A}, \hat{b}, \hat{a}_0, \hat{b}_0$ depend on $(a_l, b_l, l \ge 0)$ and r depends on ϵ

- More conservative but simpler to solve
- A feasible, or optimal, x for SOCP always satisfies the chance constraint
- Feasible set of safe approximation is inner approximation of feasible set of chance constrained problem

Safe approximation
Derivation

Derivation of inner approximation of CCP feasible set relies on

1. Concentration inequalities
    - Upper bound tail probability (violation probability of chance constraint)
    - … in terms of distribution properties, e.g., variance, log moment generating function ψ_Y

2. sub-Gaussian random variables
    - Upper bound distribution properties (e.g. ψ_Y) of uncertain parameters ζ by known distribution properties, e.g., those of Gaussian random variable

We explain each in turn

Concentration inequalities

Markov's inequality

Let Y be a nonnegative random variable with finite mean $EY < \infty$

$$\mathbb{P}\left(Y \geq t\right) \;\leq\; \frac{EY}{t}$$

Proof: for $t > 0$, take expectation on $Y/t \geq \delta(Y \geq t)$ indicator function

Concentration inequalities

Markov's inequality

Let Y be a nonnegative random variable with finite mean $EY < \infty$

$$\mathbb{P}\left(Y \geq t\right) \;\leq\; \frac{EY}{t}$$

Proof: for $t > 0$, take expectation on $Y/t \geq \delta(Y \geq t)$ indicator function

For any nonnegative and nondecreasing function ϕ

$$\mathbb{P}\left(Y \geq t\right) \;\leq\; \frac{E(\phi(Y))}{\phi(t)}$$

Proof: $\delta(Y \geq t) = \delta(\phi(Y) \geq \phi(t))$

Concentration inequalities

Chebyshev's inequality

Let X be a random variable with finite variance var$(X) < \infty$

$$\mathbb{P}\left(|X - EX| \geq t \right) \leq \frac{\text{var}(X)}{t^2}$$

Proof: take $\phi(t) := t^2$ in Markov's inequality

Concentration inequalities

Chebyshev's inequality

Let X be a random variable with finite variance $\text{var}(X) < \infty$

$$\mathbb{P}\left(|X - EX| \geq t\right) \leq \frac{\text{var}(X)}{t^2}$$

Proof: take $\phi(t) := t^2$ in Markov's inequality

For independent random variables $X_1, \ldots, X_n$ with finite variances $\text{var}(X_i) < \infty$

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_i (X_i - EX_i)\right| \geq t\right) \leq \frac{\sum_i \text{var}(X_i)}{n^2 t^2} = \frac{\sigma_n^2}{n t^2}$$

where $\sigma_n^2 := n^{-1} \sum_i \text{var}(X_i)$

Concentration inequalities

Chernoff bound

Let Y be a random variable with finite mean $EY < \infty$

$E(e^{\lambda Y})$ is moment-generating function of Y.

Define log moment-generating function:

$$\psi_Y(\lambda) \;:=\; \ln E\left(e^{\lambda Y}\right), \qquad \lambda \in \mathbb{R}$$

and its conjugate function:

$$\psi_Y^*(t) \;:=\; \sup_{\lambda \in \mathbb{R}} \left(t\lambda - \psi_Y(\lambda)\right), \qquad t \in \mathbb{R}$$

Then $\psi_Y(0) = 0$, $\psi_Y(\lambda) \geq \lambda EY$

Concentration inequalities

Chernoff bound

Let Y be a random variable with finite mean $EY < \infty$

Three equivalent forms of Chernoff bound:

1. For $t \geq EY$

$$\mathbb{P}(Y \geq t) \ \leq \ e^{-\psi_Y^*(t)}$$

Proof: take $\phi(t) := e^{\lambda t}$ which is nonnegative and nondecreasing for $\lambda \geq 0$

2. For $t \in \mathbb{R}$

$$\mathbb{P}(Y \geq t) \ \leq \ \exp\left(- \sup_{\lambda \geq 0} \left(t\lambda - \psi_Y(\lambda) \right) \right)$$

3. For $t \in \mathbb{R}$

$$\ln \mathbb{P}(Y \geq t) \ \leq \ \inf_{\lambda \geq 0} \ \ln \left(e^{-\lambda t} E e^{\lambda Y} \right)$$

Concentration inequalities
Chernoff bound

Let $Y := \frac{1}{n}\sum_i X_i$ be sample mean of independent random variables X_i with $EX_i < \infty$, $i = 1,\ldots,n$

1. If X_i are independent, then $\psi_Y(\lambda) = \sum_i \psi_{X_i}(\lambda/n)$ and

$$\psi_Y^*(t) = \sup_{\lambda\in\mathbb{R}} \sum_i \left(t\lambda - \psi_{X_i}(\lambda)\right) \leq \sum_i \psi_{X_i}^*(t)$$

with "=" if X_i are iid

Concentration inequalities

Chernoff bound

Let $Y := \frac{1}{n}\sum_i X_i$ be sample mean of independent random variables X_i with $EX_i < \infty$, $i = 1,\ldots,n$

1. If X_i are independent, then $\psi_Y(\lambda) = \sum_i \psi_{X_i}(\lambda/n)$ and

$$\psi_Y^*(t) = \sup_{\lambda\in\mathbb{R}} \sum_i \left(t\lambda - \psi_{X_i}(\lambda)\right) \le \sum_i \psi_{X_i}^*(t) \qquad \text{with "=" if } X_i \text{ are iid}$$

$$\mathbb{P}\left(\frac{1}{n}\sum_i X_i \ge t\right) \le e^{-\psi_Y^*(t)} = e^{-nI_n(t)}, \qquad t \ge \frac{1}{n}\sum_i EX_i$$

where $I_n(t)$ is called a rate function defined as

$$I_n(t) := \sup_{\lambda\in\mathbb{R}}\left(t\lambda - \frac{1}{n}\sum_i \psi_{X_i}(\lambda)\right), \qquad t \ge \frac{1}{n}\sum_i EX_i$$

Concentration inequalities

Chernoff bound

Let $Y := \frac{1}{n}\sum_i X_i$ be sample mean of independent random variables X_i with $EX_i < \infty$, $i = 1,\ldots,n$

1. If X_i are independent, then $\psi_Y(\lambda) = \sum_i \psi_{X_i}(\lambda/n)$ and

$$\psi_Y^*(t) = \sup_{\lambda\in\mathbb{R}} \sum_i \left(t\lambda - \psi_{X_i}(\lambda)\right) \le \sum_i \psi_{X_i}^*(t) \qquad \text{with "=" if } X_i \text{ are iid}$$

$$\mathbb{P}\left(\frac{1}{n}\sum_i X_i \ge t\right) \le e^{-\psi_Y^*(t)} = e^{-nI_n(t)}, \qquad t \ge \frac{1}{n}\sum_i EX_i$$

where $I_n(t)$ is called a rate function defined as

$$I_n(t) := \sup_{\lambda\in\mathbb{R}}\left(t\lambda - \frac{1}{n}\sum_i \psi_{X_i}(\lambda)\right) \le \frac{1}{n}\sum_i \psi_{X_i}^*(t) \quad \text{with "=" if } X_i \text{ are iid}$$

Concentration inequalities

Chernoff bound

Let $Y := \frac{1}{n} \sum_i X_i$ be sample mean of independent random variables X_i with $EX_i < \infty$, $i = 1, \ldots, n$

2. If X_i are iid

$$\mathbb{P}\left(\frac{1}{n} \sum_i X_i \geq t \right) \leq e^{-n\psi^*_{X_1}(t)} \qquad t \geq EX_1$$

Gaussian random variable

Let Y be Gaussian random variable with $\mu := EY$ and standard deviation $\sigma := \sqrt{\text{var}(Y)}$

Log moment-generating function:

$$\psi_{\text{G}}(\lambda) \;:=\; \ln E\left(e^{\lambda Y}\right) \;=\; \mu\lambda + \frac{\sigma^2}{2}\lambda^2, \qquad \lambda \in \mathbb{R}$$

and its conjugate function:

$$\psi^*_{\text{G}}(t) \;:=\; \sup_{\lambda\in\mathbb{R}} \left(t\lambda - \psi_Y(\lambda)\right) \;=\; \frac{(t-\mu)^2}{2\sigma^2}, \qquad t \in \mathbb{R}$$

Chernoff bound for Gaussian random var:

$$\mathbb{P}(Y > \mu + r\sigma) \;\leq\; e^{-r^2/2}, \qquad r \geq 0$$

probability of Gaussian r.v. exceeding r std above its mean decays exponentially in r^2

Gaussian random variable

Weighted sum of independent Gaussians

Let $Y := \sum_i a_i X_i$ of independent Gaussian r.v. X_i with $\left(\mu_i, \sigma_i^2\right)$

Then $Y \sim N\left(\sum_i a_i \mu_i, \sum_i a_i^2 \sigma_i^2\right)$. Hence

$$\psi_Y(\lambda) = \ln E e^{\lambda Y} = \lambda \sum_i a_i \mu_i + \frac{\lambda^2}{2} \sum_i a_i^2 \sigma_i^2, \qquad \lambda \in \mathbb{R}$$

$$\psi_Y^*(t) = \sup_{\lambda \in \mathbb{R}} \left(t\lambda - \phi_Y(\lambda)\right) = \frac{(t - \sum_i a_i \mu_i)^2}{2 \sum_i a_i^2 \sigma_i^2}, \qquad t \in \mathbb{R}$$

$$\mathbb{P}\left(\sum_i a_i (X_i - \mu_i) > r \sqrt{\sum_i a_i^2 \sigma_i^2}\right) \leq e^{-r^2/2}, \qquad r \geq 0$$

Gaussian random variable

Sample mean

Let $Y := \frac{1}{n}\sum_i X_i$ be the sample mean of independent Gaussian r.v. X_i with $\left(\mu_i, \sigma_i^2\right)$

Then $Y \sim N\left(\frac{1}{n}\sum_i \mu_i, \frac{1}{n}v_n\right)$ where $v_n := \frac{1}{n}\sum_i \sigma_i^2$ is avg var. Hence

$$\mathbb{P}\left(\frac{1}{n}\sum_i (X_i - \mu_i) > t\right) \;\leq\; e^{-nt^2/2v_n}, \qquad t \geq 0$$

If X_i are iid then

$$\mathbb{P}\left(\frac{1}{n}\sum_i X_i - \mu_1 > t\right) \;\leq\; e^{-nt^2/2\sigma_1^2}, \qquad t \geq 0$$

sub-Gaussian random variable

A r.v. Y is sub-Gaussian with (μ, σ^2) if its log moment-generating function is upper bounded by that of the Gaussian r.v.:

$$\psi_Y(\lambda) \;\leq\; \psi_{\mathsf{G}}(\lambda) \;=\; \mu\lambda + \frac{\sigma^2}{2}\lambda^2, \qquad \lambda \in \mathbb{R}$$

Hence conjugate function:

$$\psi_Y^*(t) \;\geq\; \psi_{\mathsf{G}}^*(t) \;=\; \frac{(t-\mu)^2}{2\sigma^2}, \qquad t \in \mathbb{R}$$

Chernoff bound:

$$\mathbb{P}(Y > t) \;\leq\; e^{-\psi_Y^*(t)} \;\leq\; e^{-(t-\mu)^2/2\sigma^2}, \qquad t \geq EY$$

Tail probability of sub-Gaussian r.v. decays more rapidly than that of the bounding Gaussian r.v.
As far as Chernoff bound is concern, sub-Gaussian r.v. behaves like its bounding Gaussian r.v.

sub-Gaussian random variable

Weighted sum of independent sub-Gaussians

Let $Y := \sum_i a_i X_i$ of independent sub-Gaussian r.v. X_i with $\left(\mu_i, \sigma_i^2\right)$

$$\phi_{X_i}(\lambda) \;\leq\; \mu_i \lambda + \frac{\sigma_i^2}{2}\lambda^2,$$

Then Y is sub-Gaussian with $(\mu, \sigma^2) := \left(\sum_i a_i \mu_i, \sum_i a_i^2 \sigma_i^2\right)$:

$$\psi_Y(\lambda) \;\leq\; \mu\lambda + \frac{\sigma^2}{2}\lambda^2$$

$$\mathbb{P}\left(Y \geq t\right) \;\leq\; \exp\left(-\frac{\left(t-\mu\right)^2}{2\sigma^2}\right), \qquad t \geq EY$$

Chernoff bound of sub-Gaussian weighted sum is same as that of bounding Gaussian weighted sum

Chance constrained optimization

$$\min_{x \in X} \quad c(x) \qquad \text{s.t.} \qquad F_\zeta(h(x)) \geq p$$

Will introduce two techniques to deal with chance constrained opt

1. Tractable instances
   - … when constraint functions h_i and probability measure $\mathbb{P}$ have certain concavity properties
   - Study conditions for feasible set to be convex and for strong duality and dual optimality

2. Safe approximation through concentration inequalities
   - Safe approximation: more conservative but simpler to solve
   - Upper bounding violation probability using concentration inequality (e.g. Chernoff bound)
   - Upper bounding distribution of ζ by known distribution (e.g. sub-Gaussian)

Safe approximation

Chance constrained LP

Consider

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(\sum_{l=1}^{k} \left(a_l^\top x - b_l\right) \zeta_l \leq -\left(a_0^\top x - b_0\right) \right) \; \geq \; 1 - \epsilon$$

where ζ_l are independent sub-Gaussian with (μ_i, σ_i^2):

$$\psi_{\zeta_l}(\lambda) \; \leq \; \mu_l \lambda + \frac{\sigma_l^2}{2} \lambda^2, \qquad\qquad \lambda \in \mathbb{R}$$

An optimization problem is a safe approximation of the chance constrained LP if feasible set of the safe approximation is a subset (inner approximation) of feasible set of the chance constrained LP

$\Longrightarrow$ an optimal solution of safe approximation satisfies the chance constraint

Safe approximation

Chance constrained LP

Consider

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(\sum_{l=1}^{k} (a_l^\top x - b_l)\, \zeta_l \le -(a_0^\top x - b_0) \right) \ \ge\ 1 - \epsilon$$

where ζ_l are independent sub-Gaussian with (μ_l, σ_l^2):

$$\psi_{\zeta_l}(\lambda) \ \le\ \mu_l \lambda + \frac{\sigma_l^2}{2}\lambda^2, \qquad \lambda \in \mathbb{R}$$

Let $A^\top \ := \ [a_1 \ \cdots \ a_k]$ and $b := (b_1, \ldots, b_k)$. The chance constrained LP is:

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(\zeta^\top (Ax - b) \le -(a_0^\top x - b_0) \right) \ \ge\ 1 - \epsilon$$

Safe approximation

Chance constrained LP

Consider

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(\zeta^\top (Ax - b) \leq -(a_0^\top x - b_0)\right) \;\geq\; 1 - \epsilon$$

Theorem

The following SOCP is a safe approximation:

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad r\|\sqrt{\Sigma}(Ax - b)\|_2 \leq -(\hat{a}_0^\top x - \hat{b}_0)$$

where $r := \sqrt{2\ln(1/\epsilon)}$ and

$$\hat{a}_0 := \; a_0 + A^\top \mu \in \mathbb{R}^n, \qquad \hat{b}_0 := \; b_0 + b^\top \mu \in \mathbb{R}$$

$$\mu := \; (\mu_1, \ldots, \mu_k), \qquad \Sigma := \; \text{diag}\left(\sigma_1^2, \ldots, \sigma_k^2\right)$$

Safe approximation

Proof

Fix $x \in \mathbb{R}^n$. Let $c_l(x) := a_l^\top x - b_l$, $l = 0, \ldots, k$ and $Y(x) := \sum_{l=1}^{k} c_l(x)\zeta_l$

Violation probability: $\mathbb{P}\left(Y(x) > -c_0(x)\right)$

Safe approximation

Proof

Fix $x \in \mathbb{R}^n$. Let $c_l(x) := a_l^\top x - b_l,\ l = 0, \ldots, k$ and $Y(x) := \sum_{l=1}^{k} c_l(x)\zeta_l$

Violation probability: $\mathbb{P}\left(Y(x) > -c_0(x)\right)$ and $Y(x)$ is sub-Gaussian with

$$\left(\mu(x), \sigma^2(x)\right) \ := \ \left(\sum_l c_l(x)\mu_l, \sum_l c_l^2(x)\sigma_l^2\right)$$

i.e.

$$\psi_{Y(x)}(\lambda) \ \leq \ \mu(x)\lambda + \frac{\sigma^2(x)}{2}\lambda^2$$

Safe approximation

Proof

Fix $x \in \mathbb{R}^n$. Let $c_l(x) := a_l^\top x - b_l,\ l = 0,\ldots,k$ and $Y(x) := \sum_{l=1}^{k} c_l(x)\zeta_l$

Violation probability: $\mathbb{P}\left(Y(x) > -c_0(x)\right)$ and $Y(x)$ is sub-Gaussian with

$$\left(\mu(x), \sigma^2(x)\right) \ := \ \left(\sum_l c_l(x)\mu_l, \sum_l c_l^2(x)\sigma_l^2\right)$$

i.e.

$$\psi_{Y(x)}(\lambda) \ \le \ \mu(x)\lambda + \frac{\sigma^2(x)}{2}\lambda^2$$

Hence Chernoff bound on $Y(x)$ is:

$$\ln \mathbb{P}\left(Y(x) > -c_0(x)\right) \ \le \ \inf_{\lambda \ge 0} \ \psi_{Y(x)}(\lambda) + c_0(x)\lambda \ \le \ \inf_{\lambda \ge 0} \ (c_0(x) + \mu(x))\lambda + \frac{\sigma^2(x)}{2}\lambda^2$$

Safe approximation

Proof

Fix $x \in \mathbb{R}^n$. Let $c_l(x) := a_l^\top x - b_l,\ l = 0,\ldots,k$ and $Y(x) := \sum_{l=1}^{k} c_l(x)\zeta_l$

The minimum is attained at $\lambda(x) := \left[-(c_0(x) + \mu(x))/\sigma^2(x)\right]^+$ and hence

$$\ln \mathbb{P}\left(Y(x) > -c_0(x)\right) \leq -\frac{(c_0(x) + \mu(x))^2}{2\sigma^2(x)}$$

Safe approximation

Proof

Fix $x \in \mathbb{R}^n$. Let $c_l(x) := a_l^\top x - b_l,\ l = 0,\ldots,k$ and $Y(x) := \sum_{l=1}^{k} c_l(x)\zeta_l$

The minimum is attained at $\lambda(x) := \left[-(c_0(x) + \mu(x))/\sigma^2(x)\right]^+$ and hence

$$\ln \mathbb{P}\left(Y(x) > -c_0(x)\right) \le -\frac{(c_0(x) + \mu(x))^2}{2\sigma^2(x)}$$

Hence x is feasible if

$$-\frac{(c_0(x) + \mu(x))^2}{2\sigma^2(x)} \le \ln \epsilon \quad \Longleftrightarrow \quad \sqrt{2\ln(1/\epsilon)}\,\sigma(x) \le -(c_0(x) + \mu(x))$$

Safe approximation

Proof

Fix $x \in \mathbb{R}^n$. Let $c_l(x) := a_l^\top x - b_l,\ l = 0,\ldots,k$ and $Y(x) := \sum_{l=1}^{k} c_l(x)\zeta_l$

The minimum is attained at $\lambda(x) := \left[-(c_0(x) + \mu(x))/\sigma^2(x)\right]^+$ and hence

$$\ln \mathbb{P}\left(Y(x) > -c_0(x)\right) \leq -\frac{(c_0(x) + \mu(x))^2}{2\sigma^2(x)}$$

Hence x is feasible if

$$-\frac{(c_0(x) + \mu(x))^2}{2\sigma^2(x)} \leq \ln \epsilon \quad \Longleftrightarrow \quad \sqrt{2\ln(1/\epsilon)}\sigma(x) \leq -(c_0(x) + \mu(x))$$

or if

$$\sqrt{2\ln(1/\epsilon)}\sqrt{\sum_l \sigma_l^2 c_l^2(x)} \leq -\left(c_0(x) + \sum_l \mu_l c_l(x)\right) \Leftrightarrow r\|\sqrt{\Sigma}(Ax - b)\|_2 \leq -(\hat{a}_0^\top x - \hat{b}_0)$$

Comparison: uncertain LPs

Example

Consider uncertain LP

$$\min_{x \in \mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad (a_0 + a_1\zeta_1 + a_2\zeta_2)^\mathsf{T} x \ \leq \ 0$$

where uncertain parameter $\zeta := (\zeta_1, \zeta_2)$ takes value in $Z_\infty := \{\zeta : \|\zeta\|_\infty \leq 1\}$

Comparison: uncertain LPs

Example

Consider uncertain LP

$$\min_{x \in \mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad (a_0 + a_1\zeta_1 + a_2\zeta_2)^\mathsf{T} x \ \leq \ 0$$

where uncertain parameter $\zeta := (\zeta_1, \zeta_2)$ takes value in $Z_\infty := \{\zeta : \|\zeta\|_\infty \leq 1\}$

1. Robust counterpart:

$$\min_{x \in \mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad a_0^\mathsf{T} x + \max_{\zeta \in Z_\infty} \left(a_1\zeta_1 + a_2\zeta_2\right)^\mathsf{T} x \ \leq \ 0$$

Comparison: uncertain LPs

Example

Consider uncertain LP

$$\min_{x \in \mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad (a_0 + a_1\zeta_1 + a_2\zeta_2)^\mathsf{T} x \;\leq\; 0$$

where uncertain parameter $\zeta := (\zeta_1, \zeta_2)$ takes value in $Z_\infty := \{\zeta : \|\zeta\|_\infty \leq 1\}$

1. Robust counterpart:

$$\min_{x \in \mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad a_0^\mathsf{T} x + \max_{\zeta \in Z_\infty} \left(a_1\zeta_1 + a_2\zeta_2\right)^\mathsf{T} x \;\leq\; 0$$

which is equivalent to LP: $\min_{x \in \mathbb{R}^n} \; c^\mathsf{T} x \quad \text{s.t.} \quad x \in X_1$ where (solving max in closed form)

$$X_1 := \left\{ x \in \mathbb{R}^n : a_0^\mathsf{T} x + \hat{A} x \leq 0 \right\} \qquad \text{with} \qquad \hat{A} := \begin{bmatrix} (+a_1 + a_2)^\mathsf{T} \\ (+a_1 - a_2)^\mathsf{T} \\ (-a_1 + a_2)^\mathsf{T} \\ (-a_1 - a_2)^\mathsf{T} \end{bmatrix}$$

Comparison: uncertain LPs

Example

Consider uncertain LP

$$\min_{x\in\mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad (a_0 + a_1\zeta_1 + a_2\zeta_2)^\top x \ \le \ 0$$

where uncertain parameter $\zeta := (\zeta_1, \zeta_2)$ takes value in $Z_\infty := \{\zeta : \|\zeta\|_\infty \le 1\}$

2. Chance constrained formulation:

$$\min_{x\in\mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad \mathbb{P}\left((a_0 + a_1\zeta_1 + a_2\zeta_2)^\top x \le 0\right) \ \ge \ 1 - \epsilon$$

Denote its feasible set by X_2

Comparison: uncertain LPs

Example

Consider uncertain LP

$$\min_{x\in\mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad (a_0 + a_1\zeta_1 + a_2\zeta_2)^\mathsf{T} x \ \leq \ 0$$

where uncertain parameter $\zeta := (\zeta_1, \zeta_2)$ takes value in $Z_\infty := \{\zeta : \|\zeta\|_\infty \leq 1\}$

3. Safe approximation: Suppose ζ_l are independent and zero-mean r.v. Since they take values in $[-1,1]$, they are sub-Gaussian with $(\mu_l, \sigma_l^2) = (0,1)$ (Hoeffinding's Lemma)

Comparison: uncertain LPs

Example

Consider uncertain LP

$$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad (a_0 + a_1\zeta_1 + a_2\zeta_2)^\top x \ \leq \ 0$$

where uncertain parameter $\zeta := (\zeta_1, \zeta_2)$ takes value in $Z_\infty := \{\zeta : \|\zeta\|_\infty \leq 1\}$

3. Safe approximation: Suppose ζ_l are independent and zero-mean r.v. Since they take values in $[-1,1]$, they are sub-Gaussian with $(\mu_l, \sigma_l^2) = (0,1)$ (Hoeffinding's Lemma)

   Therefore the SOCP is a safe approximation:

   $$\min_{x \in \mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad a_0^\top x \ + \ r\|Ax\|_2 \leq 0$$

   where $r := \sqrt{2\ln(1/\epsilon)}$, $A := [a_1 \ a_2]^\top$

   Feasible set is $X_3 \ := \ \left\{ x \in \mathbb{R}^n : \begin{bmatrix} A \\ -(1/r)a_0^\top \end{bmatrix} x \in K_{\text{soc}} \right\}$

Comparison: uncertain LPs

Example

Consider uncertain LP

$$\min_{x\in\mathbb{R}^n} \quad c^\mathsf{T} x \quad \text{s.t.} \quad (a_0 + a_1\zeta_1 + a_2\zeta_2)^\mathsf{T} x \ \leq \ 0$$

where uncertain parameter $\zeta := (\zeta_1, \zeta_2)$ takes value in $Z_\infty := \{\zeta : \|\zeta\|_\infty \leq 1\}$

- Feasible sets X_1, X_3 are convex, X_2 of chance constrained opt may not.
- $X_1 \subseteq X_2$, $X_3 \subseteq X_2$
- But neither X_1 nor X_3 may contain the other, depending on ϵ, i.e., robust LP may not be more conservative than safe approximation of chance constrained LP

Comparison: uncertain LPs

Example

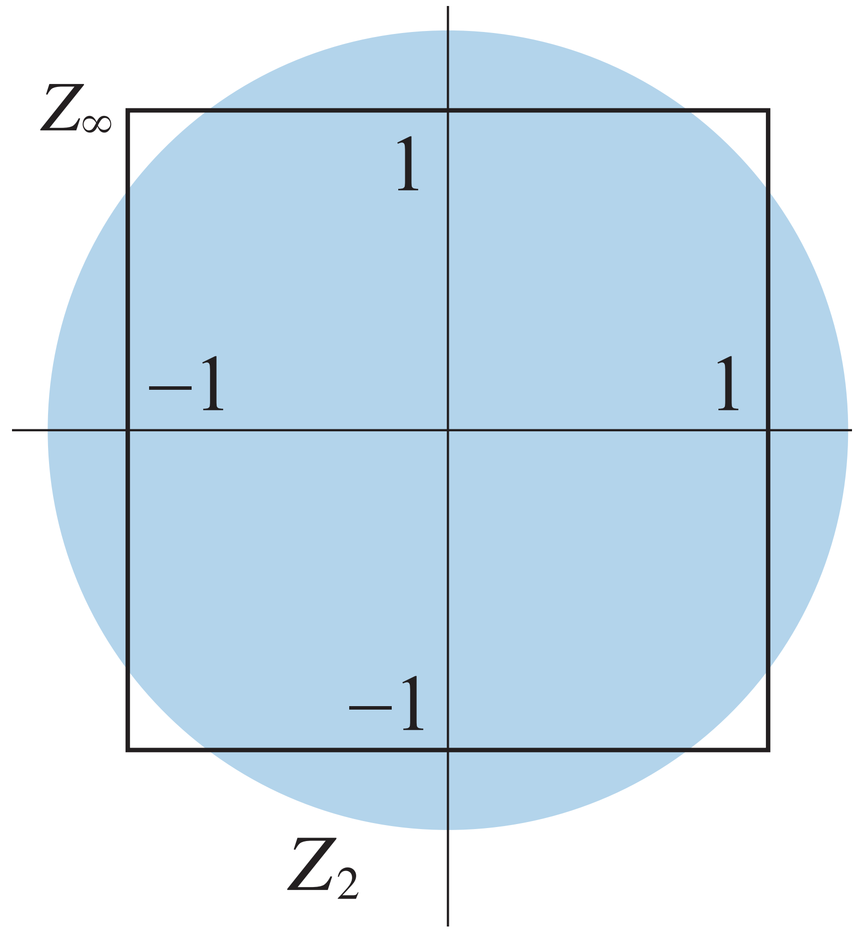
- This is because safe approximation (SOCP) is equivalent to the robust LP:

$$\min_{x\in\mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad a_0^\top x + \max_{\zeta\in Z_2} \left(a_1\zeta_1 + a_2\zeta_2\right)^\top x \ \leq \ 0$$

where $Z_2 := \{\zeta \in \mathbb{R}^2 : \|\zeta\|_2 \leq \sqrt{2\ln(1/\epsilon)}\}$

Compare with robust LP:

$$\min_{x\in\mathbb{R}^n} \quad c^\top x \quad \text{s.t.} \quad a_0^\top x + \max_{\zeta\in Z_\infty} \left(a_1\zeta_1 + a_2\zeta_2\right)^\top x \ \leq \ 0$$



Summary
Concentration inequalities

	Inequality	Assumptions
Markov's	$\mathbb{P}(Y \geq t) \leq \frac{E(\phi(Y))}{\phi(t)}$	$\phi(Y) \geq 0$, $\phi(t) > 0$, $EY < \infty$
Chebyshev's	$\mathbb{P}(\lvert X - EX\rvert \geq t) \leq \mathrm{var}(X)/t^2$ $\mathbb{P}\left(\left\lvert\frac{1}{n}\sum_i (X_i - EX_i)\right\rvert \geq t\right) \leq \frac{(1/n)\sum_i \mathrm{var}(X_i)}{nt^2}$	$\mathrm{var}(X) < \infty$, $t > 0$ $\mathrm{var}(X_i) < \infty$, independent X_i, $t > 0$
Chernoff	$\mathbb{P}(Y \geq t) \leq e^{-\psi_Y^*(t)}$ $\mathbb{P}(Y \geq t) \leq \exp\left(-\sup_{\lambda \geq 0}(t\lambda - \psi_Y(\lambda))\right)$ $\mathbb{P}\left(\frac{1}{n}\sum_i X_i \geq t\right) \leq e^{-n\psi_{X_1}^*(t)}$	$EY < \infty$, $t \geq EY$ $EY < \infty$, $t \in \mathbb{R}$ iid X_i, $EX_i < \infty$, $t \geq E(X_1)$
sub-Gaussian	$\mathbb{P}(Y \geq t) \leq e^{-(t-\mu)^2/2\sigma^2}$ $\mathbb{P}(\sum_i a_i X_i \geq t) \leq \exp\left(-\frac{(t - \sum_i a_i \mu_i)^2}{2\sum_i a_i^2 \sigma_i^2}\right)$ $\mathbb{P}\left(\max_{i=1}^n X_i \geq t\right) \leq \sigma\sqrt{2\ln n}/t$	sub-Gaussian Y, $EY < \infty$, $t \geq EY$ indep. sub-Gaussian X_i, $EX_i < \infty$, $t \geq EY$ sub-Gaussian X_i, $t > 0$
Hoeffding's lemma	$\psi_Y(\lambda) \leq (1/8)(b-a)^2\lambda^2$	$EY = 0$, $Y \in [a, b]$ a.s.
Azuma-Hoeffding	$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \geq t\right) \leq \exp\left(-\frac{2n^2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$ $\mathbb{P}(X_n - X_0 \geq t) \leq \exp\left(-t^2/2\sum_{i=1}^n \sigma_i^2\right)$	independent zero-mean $X_i \in [a_i, b_i]$, $t \geq 0$ martingale X_i, $\lvert X_i - X_{i-1}\rvert \leq \sigma_i$, $t \geq 0$

Outline

1. Robust optimization
2. Chance constrained optimization
3. Convex scenario optimization
    - Violation probability
    - Sample complexity
    - Optimality guarantee
4. Stochastic optimization with recourse

Convex scenario opt

Consider

$$\text{RCP}: \quad c^*_{\text{RCP}} := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta) \le 0,\ \zeta \in Z \subseteq \mathbb{R}^k$$

$$\text{CCP}(\epsilon): \quad c^*_{\text{CCP}}(\epsilon) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(h(x,\zeta) \le 0\right) \ \ge\ 1-\epsilon$$

$$\text{CSP}(N): \quad c^*_{\text{CSP}}(N) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta^i) \le 0,\ i = 1,\ldots,N$$

- X : nonempty closed convex set
- $h : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$: convex (and hence continuous) in x for every uncertain parameter $\zeta \in Z$
- $\mathbb{P}$: probability measure on some probability space; $\epsilon \in [0,1]$
- $\left(\zeta^1,\ldots,\zeta^N\right)$: independent random samples each according to $\mathbb{P}$
- Linear cost: does not lose generality (can convert nonlinear cost $\min_x f(x)$ to linear cost $\min_{x,t} t$ with additional constraint $f(x) \le t$)

Convex scenario opt

Consider

$$\text{RCP}: \quad c^*_{\text{RCP}} := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta) \le 0,\ \zeta \in Z \subseteq \mathbb{R}^k$$

$$\text{CCP}(\epsilon): \quad c^*_{\text{CCP}}(\epsilon) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(h(x,\zeta) \le 0\right) \ \ge\ 1-\epsilon$$

$$\text{CSP}(N): \quad c^*_{\text{CSP}}(N) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta^i) \le 0,\ i = 1,\ldots,N$$

- RCP : deterministic, semi-infinite, generally computational hard, conservative (safe)
- CCP(ϵ) : deterministic, generally computationally hard, less conservative, need $\mathbb{P}$
- CSP(N) : randomized, finite convex program for each realization of $\zeta := \left(\zeta^1, \ldots, \zeta^N\right)$, less conservative, only need samples under $\mathbb{P}$ (not necessarily $\mathbb{P}$ itself), much more practical

Convex scenario opt

Consider

$$
\begin{aligned}
\text{RCP}: \quad & c^*_{\text{RCP}} &:= \min_{x \in X \subseteq \mathbb{R}^n} \; c^\top x \quad \text{s.t.} \quad & h(x,\zeta) \le 0,\ \zeta \in Z \subseteq \mathbb{R}^k \\
\text{CCP}(\epsilon): \quad & c^*_{\text{CCP}}(\epsilon) &:= \min_{x \in X \subseteq \mathbb{R}^n} \; c^\top x \quad \text{s.t.} \quad & \mathbb{P}\left(h(x,\zeta) \le 0\right) \;\ge\; 1-\epsilon \\
\text{CSP}(N): \quad & c^*_{\text{CSP}}(N) &:= \min_{x \in X \subseteq \mathbb{R}^n} \; c^\top x \quad \text{s.t.} \quad & h(x,\zeta^i) \le 0,\ i = 1,\ldots,N
\end{aligned}
$$

Study 3 questions on CSP(N):

- Violation probability : how likely is the random solution x^*_N of CSP(N) feasible for CCP(ϵ)?
- Sample complexity : what is min N for x^*_N to be feasible for CCP(ϵ) in expectation or probability?
- Optimality guarantee : how close is the min cost $c^*_{\text{CSP}}(N)$ to the min costs $c^*_{\text{CCP}}(\epsilon)$ and c^*_{RCP} ?

Assumption

Let $X_\zeta := \{x \in X \subseteq \mathbb{R}^n : h(x, \zeta) \leq 0\}$

$$\text{CSP}(N): \qquad c^*_{\text{CSP}}(N) := \min_{x \in X \subseteq \mathbb{R}^n} \; c^\mathsf{T} x \quad \text{s.t.} \quad h(x, \zeta^i) \leq 0,\ i = 1, \ldots, N$$

Assumption 1

- For each $\zeta \in Z, h(x, \zeta)$ is convex and continuous in x so that X_ζ is a closed convex set
- For each integer $N \geq n$ and each realization of $\zeta := \left(\zeta^1, \ldots, \zeta^N\right)$, feasible set of CSP($N$) has a nonempty interior. Moreover CSP(N) has a unique optimal solution x_N^* (can be relaxed)

Violation probability

Definition

Let $X_\zeta := \{x \in X \subseteq \mathbb{R}^n : h(x,\zeta) \leq 0\}$

Violation probability: $V(x) := \mathbb{P}\left(\left\{\zeta \in Z : x \notin X_\zeta\right\}\right)$

- For fixed $x \in X$, $V(x)$ is a deterministic value in $[0,1]$
- CCP(ϵ) is: $c^*_{\text{CCP}}(\epsilon) \; := \; \min_{x \in X \subseteq \mathbb{R}^n} \; c^\top x \quad \text{s.t.} \quad V(x) \leq \epsilon$

- For CSP(N), optimal solution x^*_N is a random variable under product measure $\mathbb{P}^N$
- Violation probability $V\left(x^*_N\right)$ of x^*_N is therefore a random variable under $\mathbb{P}^N$, taking value in $[0,1]$
- $V\left(x^*_N\right)$ may be smaller or greater than ϵ, i.e., x^*_N may or may not be feasible for CCP(ϵ)
- Goal: derive tight upper bounds on expected value and tail probability of $V\left(x^*_N\right)$

Violation probability

Definition

Let $X_\zeta := \{x \in X \subseteq \mathbb{R}^n : h(x,\zeta) \leq 0\}$

Conditional violation probability: $V\left(x_N^*\right) := \mathbb{P}\left(\left\{\zeta \in Z : x_N^* \notin X_\zeta\right\} \middle| \left(\zeta^1, \ldots, \zeta^N\right)\right)$

- A random variable under $\mathbb{P}^N$, taking value in $[0,1]$
- Relation between r.v. $V\left(x_N^*\right)$ and the (deterministic) unconditional probability $\mathbb{P}^{N+1}\left(x_N^* \notin X_\zeta\right)$ is

$$\mathbb{P}^{N+1}\left(x_N^* \notin X_\zeta\right) = \int_{Z^N} V\left(x_N^*\right) \mathbb{P}^N\left(d\zeta^1, \ldots, d\zeta^N\right) = E^N\left(V\left(x_N^*\right)\right)$$

i.e., expected value of $V\left(x_N^*\right)$ is the unconditional probability $\mathbb{P}^{N+1}\left(x_N^* \notin X_\zeta\right)$

(This unconditional probability will be later related to support constraints)

Violation probability

Uniformly supported problem

<u>Definition</u>

Consider CSP(N)

1. A constraint X_{ζ^i} is a support constraint for CSP(N) if its removal changes the optimal solution, i.e., for every realization of $(\zeta^1, \ldots, \zeta^N) \in Z^N$, $c^\top x_N^* \neq c^\top x_{N\setminus i}^*$

2. CSP(N) is uniformly supported with $s \leq n$ support constraints if every realization of $(\zeta^1, \ldots, \zeta^N) \in Z^N$ contains exactly s support constraints (a.s.). It is fully supported if $s = n$.

- A support constraint must be active at x_N^*; the converse may not hold.
- **<u>Lemma</u>**: The number of support constraints for CSP(N) is at most n

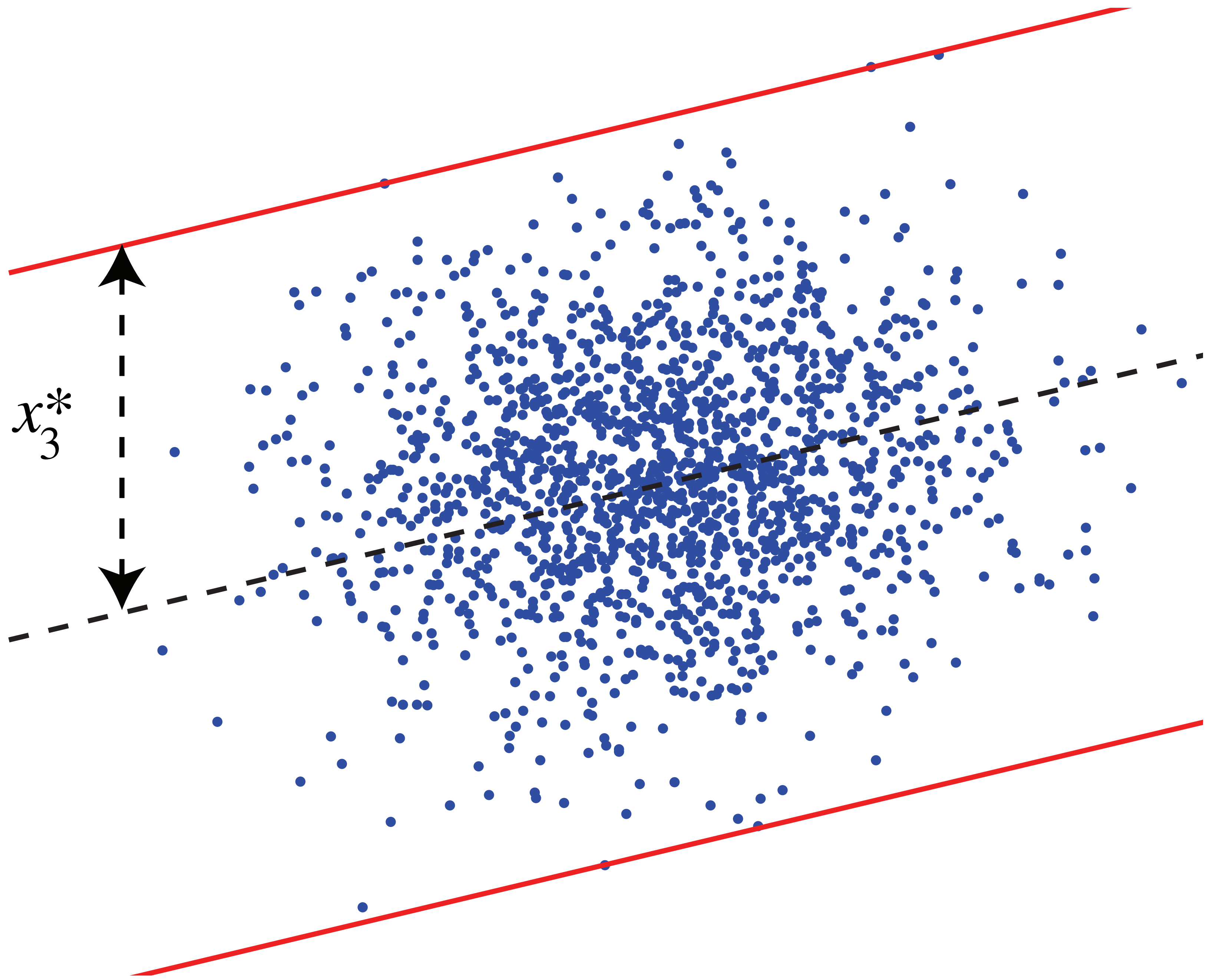
Violation probability

Uniformly supported problem

Example: fully supported problem

Construct strip of min vertical width containing all N points

$$\min_{(x_1,x_2,x_3)\in\mathbb{R}^3} \quad x_1 \quad \text{s.t.} \quad \left| b^i - (a^i x_1 + x_2) \right| \;\leq\; x_3, \quad i = 1,\ldots,N$$



For every realization of $\zeta := \big((a^i, b^i) : i = 1,\ldots,N\big)$

#support constraints $= 3 = n$

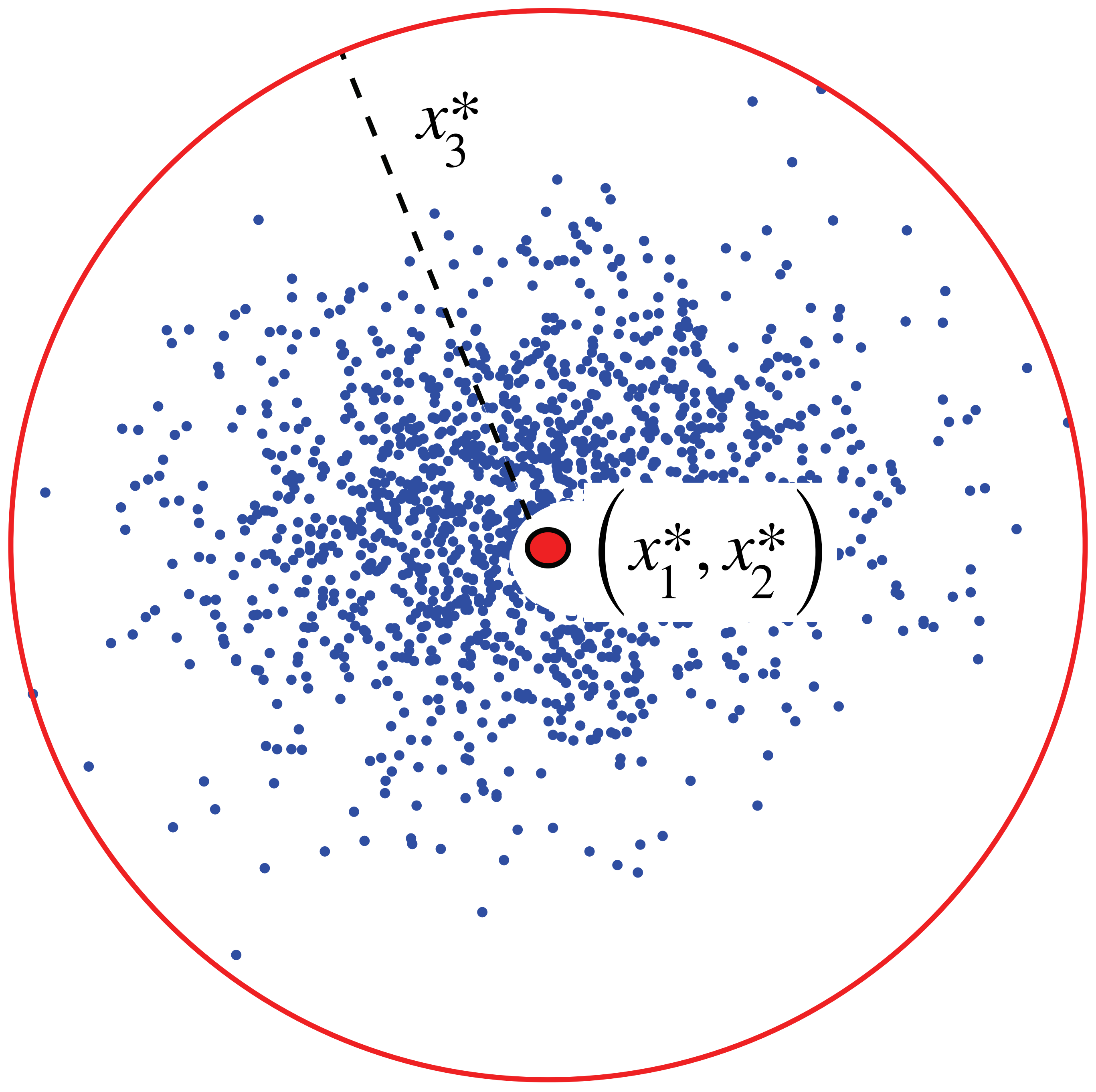
Violation probability

Uniformly supported problem

Example: uniformly supported problem

Construct circle of min radius containing all N points

$$\min_{(x_1,x_2,x_3)\in\mathbb{R}^3} \quad x_1 \quad \text{s.t.} \quad \sqrt{(a^i - x_1)^2 + (b^i - x_2)^2} \;\leq\; x_3, \quad i = 1,\ldots,N$$



For every realization of $\zeta := \left((a^i, b^i) : i = 1,\ldots,N\right)$

#support constraints $= 2 < n$

Violation probability
Expected value

Theorem [Calafiore & Campi 2005; Calafiore 2009]

Suppose Assumption 1 holds.

1. Then $E^N\left(V\left(x_N^*\right)\right) = \mathbb{P}^{N+1}\left(x_N^* \notin X_{\zeta^{N+1}}\right) \leq \dfrac{n}{N+1}$

2. If CSP$(N+1)$ is uniformly supported with $0 \leq s \leq n$ support constraints then

$$E^N\left(V\left(x_N^*\right)\right) = \mathbb{P}^{N+1}\left(x_N^* \notin X_{\zeta^{N+1}}\right) = \frac{s}{N+1}$$

- Upper bound is tight for uniformly supported problems

G. C. Calafiore and M. C. Campi, "Uncertain convex programs: Randomized solutions and confidence levels," *Math. Program.*, 2005.

G. C. Calafiore. "A note on the expected probability of constraint violation in sampled convex programs." *IEEE CCA & ISIC*, 2009.

Violation probability

Tail probability

Theorem [Campi, Garatti 2008]

Suppose Assumption 1 holds.

1. Then $\mathbb{P}^N\left(V\left(x_N^*\right) > \epsilon\right) \leq \sum_{i=0}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ Binomial tail

2. If CSP$(N+1)$ is uniformly supported with $0 \leq s \leq n$ support constraints then

$$\mathbb{P}^N\left(V\left(x_N^*\right) > \epsilon\right) = \sum_{i=0}^{s-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$$

- Upper bound is tight for uniformly supported problems

M. C. Campi and S. Garatti. "The exact feasibility of randomized solutions of uncertain convex programs." *SIAM J. Optim.*, 2008.

Violation probability

Summary

Suppose Assumption 1 holds.

1. $E^N\left(V\left(x_N^*\right)\right) \leq \dfrac{n}{N+1}$

2. $\mathbb{P}^N\left(V\left(x_N^*\right) > \epsilon\right) \leq \displaystyle\sum_{i=0}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ Binomial tail

- Binomial tail decreases rapidly as N increases
- Bounds are tight for uniformly supported problems with $0 \leq s \leq n$ support constraints
- Bounds depend only on (n, N) and ϵ.
- Not on details of cost function $c^\top x$, constraint function $h(x, \zeta)$, probability measure $\mathbb{P}$; they determine if the problem is fully supported and hence tightness of the bounds

Violation probability

Key proof idea

Partition sample space Z^N for independent samples $\left(\zeta^1, \ldots, \zeta^N\right)$ according to #support constraints

$$Z^N(I^s) := \left\{ \left(\zeta^1, \ldots, \zeta^N\right) \in Z^N : \left(X_{\zeta^i}, i \in I^s\right) \text{ are } s \text{ support constraints} \right\}$$

$$Z^N(s) := \bigcup_{I^s} Z^N(I^s) \qquad \text{conditioned on } s \text{ support constraints}$$

- $Z^N(I^s)$: vectors in Z^N whose s support constraints are indexed by $I^s \subseteq \{1, \ldots, N\}$
- $Z^N(s)$: subset of Z^N that contains exactly s support constraints

Then

$$Z^N = \bigcup_{s=0}^{n} Z^N(s) := \bigcup_{s=0}^{n} \bigcup_{I^s} Z^N(I^s)$$

Violation probability

Key proof idea

Uniformly supported with s support constraints

$$Z^N(s) := \bigcup_{I^s} Z^N(I^s), \qquad Z^N(s') = \varnothing, \quad s' \neq s$$

$$Z^N = Z^N(s) := \bigcup_{I^s} Z^N(I^s)$$

- Fully supported problem: $s = n$
- No support constraint = uniformly supported with $s = 0$ support constraint

Violation probability

Key proof idea

Uniformly supported with s support constraints

$$Z^N(s) := \bigcup_{I^s} Z^N(I^s), \qquad Z^N(s') = \varnothing, \quad s' \neq s$$

$$Z^N = Z^N(s) := \bigcup_{I^s} Z^N(I^s)$$

- Fully supported problem: $s = n$
- No support constraint = uniformly supported with $s = 0$ support constraint

Lemma [No spport constraint]

If CSP(N) has no support constraint, then $V\left(x_N^*\right) = 0$ a.s.

Hence $E^N\left(V\left(x_N^*\right)\right) = 0, \ \mathbb{P}^N\left(V\left(x_N^*\right) > \epsilon\right) = 0$

Violation probability

Key proof idea

Partition sample space Z^N for independent samples $\left(\zeta^1, \ldots, \zeta^N\right)$ according to #support constraints

$$Z^N(I^s) := \left\{ \left(\zeta^1, \ldots, \zeta^N\right) \in Z^N : \left(X_{\zeta^i}, i \in I^s\right) \text{ are } s \text{ support constraints} \right\}$$

$$Z^N(s) := \bigcup_{I^s} Z^N(I^s)$$

Lemma [Support constraints are uniformly distributed]

Suppose Assumption 1 holds. Then

$$\mathbb{P}^N\left(Z^N(I^s) \,\middle|\, Z^N(s)\right) = \left[\binom{N}{s}\right]^{-1} \qquad \forall I^s \text{ with } |I^s| = s$$

uses iid samples ζ^i

Sample complexity

Corollary

Suppose Assumption 1 holds. For any ϵ, β in $[0,1]$

1. $E^N\left(V(x_N^*)\right) \leq \beta$ if $N \geq (n/\beta) - 1$

2. $\mathbb{P}^N\left(V(x_N^*) > \epsilon\right) \leq \beta$ if $N \geq N(\epsilon, \beta)$ where

$$N(\epsilon, \beta) \; := \; \min\left\{ N \; : \; \sum_{i=0}^{n-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \; \leq \; \beta \right\}$$

Optimality guarantee

Consider

$$\text{RCP}: \quad c^*_{\text{RCP}} := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta) \leq 0,\ \zeta \in Z \subseteq \mathbb{R}^k$$

$$\text{CCP}(\epsilon): \quad c^*_{\text{CCP}}(\epsilon) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(h(x,\zeta) \leq 0\right) \geq 1-\epsilon$$

$$\text{CSP}(N): \quad c^*_{\text{CSP}}(N) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta^i) \leq 0,\ i = 1,\ldots,N$$

Study 3 questions on CSP(N):

- Violation probability : how likely is the random solution x_N^* of CSP(N) feasible for CCP(ϵ)?
- Sample complexity : what is min N for x_N^* to be feasible for CCP(ϵ) in expectation or probability?
- Optimality guarantee : how close is the min cost $c^*_{\text{CSP}}(N)$ to the min costs $c^*_{\text{CCP}}(\epsilon)$ and c^*_{RCP} ?

Optimality guarantee

Intuition

Consider

$$\text{RCP}: \quad c^*_{\text{RCP}} := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x, \zeta) \leq 0,\ \zeta \in Z \subseteq \mathbb{R}^k$$

$$\text{CCP}(\epsilon): \quad c^*_{\text{CCP}}(\epsilon) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(h(x, \zeta) \leq 0\right) \geq 1 - \epsilon$$

$$\text{CSP}(N): \quad c^*_{\text{CSP}}(N) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x, \zeta^i) \leq 0,\ i = 1, \ldots, N$$

Intuition

- Random solution x^*_N feasible for CCP(ϵ) w.h.p. connects $c^*_{\text{CSP}}(N)$ and $c^*_{\text{CCP}}(\epsilon)$
- x^*_N is however infeasible for RCP, unless $V\left(x^*_N\right) = 0$
- Key to connecting $c^*_{\text{CSP}}(N)$ and c^*_{RCP} is a perturbed RCP

Optimality guarantee

Perturbed robust program

Consider

$$\text{RCP}: \quad c^*_{\text{RCP}} := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta) \leq 0,\ \zeta \in Z \subseteq \mathbb{R}^k$$

$$\text{RCP}(\nu): \quad c^*_{\text{RCP}}(\nu) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad \bar{h}(x) := \sup_{\zeta \in Z} h(x,\zeta) \leq \nu$$

$$\text{CCP}(\epsilon): \quad c^*_{\text{CCP}}(\epsilon) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad \mathbb{P}\left(h(x,\zeta) \leq 0\right) \geq 1-\epsilon$$

$$\text{CSP}(N): \quad c^*_{\text{CSP}}(N) := \min_{x \in X \subseteq \mathbb{R}^n} c^\top x \quad \text{s.t.} \quad h(x,\zeta^i) \leq 0,\ i = 1,\ldots,N$$

- RCP = RCP(0)
- $\bar{h}(x)$ is convex in x since $h(x,\zeta)$ is convex in x for every $\zeta \in Z$

Optimality guarantee

Perturbed robust program

Definition

1. The probability of worst-case constraints is the function $p : X \times \mathbb{R}^m_+ \to [0,1]$:

$$p(x, b) \;:=\; \mathbb{P}\left(\left\{\zeta \in Z : \exists i := i(\zeta) \;\text{ s.t. }\; \bar{h}_i(x) - h_i(x, \zeta) < b_i\right\}\right)$$

2. The perturbation bound with respect to p is the function $\bar{v} : [0,1] \to \mathbb{R}^m_+$:

$$\bar{v}(\epsilon) \;:=\; \sup \left\{ b \in \mathbb{R}^m_+ : \inf_{x \in X} p(x, b) \leq \epsilon \right\}$$

where supremum is taken componentwise of vectors b

- Perturbation bound $\bar{v}(\epsilon)$ depends on constraint function $h(x, \zeta)$, uncertainty set Z, probability measure $\mathbb{P}$

Optimality guarantee
Perturbed robust program

Definition

1. The probability of worst-case constraints is the function $p : X \times \mathbb{R}^m_+ \to [0,1]$:

$$p(x, b) \;\; := \;\; \mathbb{P}\left(\left\{ \zeta \in Z : \exists i := i(\zeta) \;\; \text{s.t.} \;\; \bar{h}_i(x) - h_i(x, \zeta) < b_i \right\} \right)$$

2. The perturbation bound with respect to p is the function $\bar{v} : [0,1] \to \mathbb{R}^m_+$:

$$\bar{v}(\epsilon) \;\; := \;\; \sup \left\{ b \in \mathbb{R}^m_+ : \inf_{x \in X} p(x, b) \leq \epsilon \right\}$$

where supremum is taken componentwise of vectors b

- For fixed x, violation probability $V(x) \leq \epsilon \;\; \Leftrightarrow \;\; p(x, \bar{h}(x)) \leq \epsilon$. Hence $V(x) \leq \epsilon \;\; \Rightarrow \;\; \bar{h}(x) \leq \bar{v}(\epsilon)$

Optimality guarantee

Perturbed robust program

Lemma [Esfahani, Sutter, Lygeros 2015]

x is feasible for CCP(ϵ) $\implies$ x is feasible for RCP$(\bar{v}(\epsilon))$

Therefore, if $N \geq N(\epsilon, \beta)$ then

$$c^*_{\text{RCP}}(\bar{v}(\epsilon)) \;\leq\; c^*_{\text{CCP}}(\epsilon) \;\lessapprox\; c^\top x^*_N \;=\; c^*_{\text{CSP}}(N) \;\leq\; c^*_{\text{RCP}}$$

Lemma

w.p. $1-\beta$
Corollary

CSP(N)
relaxation

CCP(ϵ) and CSP(N) sandwiched between RCP(v) for $v = \bar{v}(\epsilon)$ and $v = 0$

Optimality guarantee

Theorem [Esfahani, Sutter, Lygeros 2015]

Suppose Assumptions 1-4 hold (see below). Given any ϵ, β in [0, 1] and any $N \geq N(\epsilon, \beta)$:

$$\mathbb{P}^N \left(c^*_{\text{RCP}} - c^*_{\text{CSP}}(N) \in [0, C(\epsilon)] \right) \;\geq\; 1 - \beta$$

$$\mathbb{P}^N \left(c^*_{\text{CSP}}(N) - c^*_{\text{CCP}}(\epsilon) \in [0, C(\epsilon)] \right) \;\geq\; 1 - \beta$$

where confidence interval is

P. M. Esfahani, T. Sutter, and J. Lygeros, "Performance bounds for the scenario approach and an extension to a class of non-convex programs," TAC, 2015.

Optimality guarantee

Theorem [Esfahani, Sutter, Lygeros 2015]

Suppose Assumptions 1-4 hold (see below). Given any ϵ, β in [0, 1] and any $N \geq N(\epsilon, \beta)$:

$$\mathbb{P}^N\left(c^*_{\text{RCP}} - c^*_{\text{CSP}}(N) \in [0, C(\epsilon)]\right) \;\geq\; 1 - \beta$$

$$\mathbb{P}^N\left(c^*_{\text{CSP}}(N) - c^*_{\text{CCP}}(\epsilon) \in [0, C(\epsilon)]\right) \;\geq\; 1 - \beta$$

where confidence interval is

$$C(\epsilon) \;:=\; \min\left\{L_{\text{RCP}}\|\bar{v}(\epsilon)\|_2,\; \max_{x\in X} c^\top x - \min_{x \in X} c^\top x\right\}$$

$$L_{\text{RCP}} \;:=\; \frac{c^\top \bar{x} - \min_{x\in X} c^\top x}{\min_i \left(v_i^{\min} - \bar{h}_i(\bar{x})\right)} \;\;\geq\;\; 0$$

Optimality guarantee

Proof idea

Assumptions

2. $V := \left\{\bar{v}(\epsilon) \in \mathbb{R}^m_+ : 0 \le \epsilon \le 1\right\}$ is compact and convex
3. For each $v \in V$
    - $\exists$ unique primal-dual optimal $(x(v), \mu(v))$ and it is continuous at v
    - Strong duality holds at $(x(v), \mu(v))$
4. Slater condition: $\exists \bar{x} \in X$ s.t. $\bar{h}(\bar{x}) < v^{\min}$ where $v_i^{\min} := \min\{v_i : v \in V\}$

Lemma [Esfahani, Sutter, Lygeros 2015]

Suppose Assumptions 1-4 hold. $c^*_{\text{RCP}}(v)$ is Lipschitz on V, i.e., for all $v_1, v_2 \in V$:

$$\left\| c^*_{\text{RCP}}(v_1) - c^*_{\text{RCP}}(v_2) \right\|_2 \le L_{\text{RCP}} \|v_1 - v_2\|_2$$

Outline

1. Robust optimization
2. Chance constrained optimization
3. Convex scenario optimization
4. Stochastic optimization with recourse
    - Stochastic LP with fixed recourse
    - Stochastic nonlinear program

Stochastic linear program

With fixed recourse

$$
\begin{aligned}
\min_{x\in\mathbb{R}^{n_1}} \quad & f(x) \ + \ E_\zeta \left(\min_{y(\omega)\in\mathbb{R}^{n_2}} \ q^\mathsf{T}(\omega) y(\omega) \right) \\
\text{s.t.} \quad & Ax = b, \ x \in K \\
& T(\omega)x + Wy(\omega) = h(\omega), \ y(\omega) \geq 0, \qquad \forall \omega \in \Omega
\end{aligned}
$$

1st-stage problem

- Cost function $f : \mathbb{R}^{n_1} \to \mathbb{R}$ is real-valued convex, K is closed convex cone
- Parameters (f, A, b, K) are certain

2nd-stage problem: semi-infinite linear program for each ω

- Recourse action $y(\omega)$ adapts to each realized $\omega \in \Omega$
- Recourse matrix W is independent of ω (i.e., fixed recourse)
- Uncertain parameter $\zeta := \zeta(\omega) := ((q(\omega), T(\omega), h(\omega)) \in \mathbb{R}^k$
- uncertainty set $Z := \{\zeta(\omega) \in \mathbb{R}^k : \omega \in \Omega\}$

Stochastic linear program

Equivalent formulation

$$\min_{x \in \mathbb{R}^{n_1}} \quad f(x) \; + \; Q(x)$$
$$\text{s.t.} \quad Ax = b, \; x \in K$$

where

$$Q(x) \; := \; E_\zeta \left(\min_{y(\omega) \geq 0} q^\mathsf{T}(\omega) y(\omega) \quad \text{s.t.} \quad W y(\omega) \; = \; h(\omega) - T(\omega)\, x \right)$$

- $Q(x)$: recourse function (or 2nd-stage expected value function)
- $Q(x)$ can be extended real-valued function and nondifferentiable
- $Q(x) = \infty$ if second-stage problem is infeasible (e.g., day-ahead schedule leads to insufficient supply when outages occur in real time)

This will be a simple conic program, but for the recourse function $Q(x)$

Example

Generation scheduling

Schedule 2 generators with same generation capacity $[0,a]$ to meet random demand $\zeta(\omega)$

1. Slow but cheap generator must be scheduled before $\zeta(\omega)$, at level $x \in [0,a]$ at unit cost c_1
2. Fast but expensive generator can be scheduled after $\zeta(\omega)$, at level $y(\omega) := y(\zeta(\omega)) \in [0,a]$ at unit cost $c_2 > c_1$
3. Suppose $\zeta(\omega) = a + \epsilon$ with prob. p, and $\zeta(\omega) = a - \epsilon$ with prob. $1 - p$

Goal: choose $(x, y(\omega))$ to meet random demand $\zeta(\omega)$ at minimum expected cost:

$$f^* \quad := \quad \min_{x \in \mathbb{R}} \; c_1 x + Q(x) \quad \text{s.t.} \quad 0 \le x \le a$$

where $Q(x) := E_\zeta \tilde{Q}(x, \zeta)$ and

$$\tilde{Q}(x, \zeta) \quad := \quad \min_{0 \le y(\omega) \le a} \; c_2 y(\omega) \quad \text{s.t.} \quad x + y(\omega) = \zeta(\omega)$$

What is the optimal solution ?

Example

Generation scheduling

2nd-stage problem:

$$\tilde{Q}(x,\zeta) \quad := \quad \min_{0\leq y(\omega)\leq a} c_2 y(\omega) \quad \text{s.t.} \quad x + y(\omega) = \zeta(\omega)$$

Since $\zeta(\omega) = a + \epsilon$ with prob. p, and $\zeta(\omega) = a - \epsilon$ with prob. $1 - p$

$$y(a+\epsilon) = \begin{cases} a+\epsilon-x & \text{if} \quad x \geq \epsilon \\ \text{infeasible} & \text{if} \quad x < \epsilon \end{cases}, = \qquad \tilde{Q} = \begin{cases} c_2(a+\epsilon-x) & \text{if} \quad x \geq \epsilon \\ \infty & \text{if} \quad x < \epsilon \end{cases}$$

$$y(a-\epsilon) = \begin{cases} a-\epsilon-x & \text{if} \quad x \leq a-\epsilon \\ \text{infeasible} & \text{if} \quad x > a-\epsilon \end{cases} \qquad \tilde{Q} = \begin{cases} c_2(a-\epsilon-x) & \text{if} \quad x \leq a-\epsilon \\ \infty & \text{if} \quad x > a-\epsilon \end{cases}$$

If $x < \epsilon$ or $x > a - \epsilon$, then $\tilde{Q}(x,\zeta) = \infty$ with probabilities p or $1 - p$ respectively and $Q(x) = E_\zeta \tilde{Q}(x,\zeta) = \infty$. Therefore

$$C_2 \ := \ \text{dom}(Q) \ := \ \{x : \epsilon \leq x \leq a - \epsilon\}$$

Example

Generation scheduling

2nd-stage problem:

$$\tilde{Q}(x,\zeta) \quad := \quad \min_{0 \le y(\omega) \le a} \; c_2 y(\omega) \quad \text{s.t.} \quad x + y(\omega) = \zeta(\omega)$$

Suppose $\zeta(\omega) = a + \epsilon$ with prob. p, and $\zeta(\omega) = a - \epsilon$ with prob. $1 - p$

$$y(a+\epsilon) = \begin{cases} a + \epsilon - x & \text{if} \;\; x \ge \epsilon \\ \text{infeasible} & \text{if} \;\; x < \epsilon \end{cases}, = \qquad \tilde{Q} = \begin{cases} c_2(a+\epsilon-x) & \text{if} \;\; x \ge \epsilon \\ \infty & \text{if} \;\; x < \epsilon \end{cases}$$

$$y(a-\epsilon) = \begin{cases} a - \epsilon - x & \text{if} \;\; x \le a - \epsilon \\ \text{infeasible} & \text{if} \;\; x > a - \epsilon \end{cases} \qquad \tilde{Q} = \begin{cases} c_2(a-\epsilon-x) & \text{if} \;\; x \le a - \epsilon \\ \infty & \text{if} \;\; x > a - \epsilon \end{cases}$$

On C_2, $Q(x) = E_\zeta \tilde{Q}(x,\zeta)$ is affine in x

$$Q(x) \;=\; pc_2(a+\epsilon-x) + (1-p)c_2(a-\epsilon-x) \;=\; c_2(a+\epsilon(2p-1)) \,-\, c_2 x$$

Example

Generation scheduling

Therefore

$$f^* \quad := \quad \min_{x \in \mathbb{R}} \ (c_1 - c_2)x + c_2(a + \epsilon(2p - 1)) \quad \text{s.t.} \quad \epsilon \le x \le a - \epsilon$$

Solution:

Since $c_2 < c_1$, optimal solution is:

$$x^* = a - \epsilon, \qquad f^* = c_1(a - \epsilon) + 2c_2\epsilon p$$

Therefore

1. The cheap generator always produces at the lower level $a - \epsilon$ of the random demand
2. The expensive generator will pick up the slack, 2ϵ with probability p

Recourse function $Q(x)$

Lemma

Suppose the recourse is fixed (W independent of ω) and $E\zeta^2 < \infty$.

1. $Q(x)$ is convex and Lipschitz on $\text{dom}(Q) := \{x : Q(x) < \infty\}$
2. If distribution function of ζ is absolutely continuous, then $Q(x)$ is differentiable on ri(dom(Q))
3. Suppose ζ takes finitely many values. Then
   - dom(Q) is closed, convex, and polyhedral
   - $Q(x)$ is piecewise linear and convex on dom(Q)

Summary: for two-stage problem with fixed recourse, if $E\zeta^2 < \infty$, then $Q(x)$ is convex and hence subdifferentiable

Hence $\min_{x \in \mathbb{R}^{n_1}} \; f(x) + Q(x)$ s.t. $Ax = b,\ x \in K$ is **nonsmooth conic** program

Strong duality and KKT

Nonsmooth conic program: $f^* \;:=\; \min_{x \in \mathbb{R}^{n_1}} \; f(x) + Q(x) \quad \text{s.t.} \quad Ax = b,\; x \in K$

where f is convex and $K \subseteq \mathbb{R}^{n_1}$ is a closed convex cone

Dual cone: $K^* \;:=\; \{\xi \in \mathbb{R}^{n_1} : \xi^\mathsf{T} x \geq 0 \;\; \forall x \in K\}$

Lagrangian:

$$L(x, \lambda, \mu) \;:=\; f(x) + Q(x) - \lambda^\mathsf{T}(Ax - b) - \mu^\mathsf{T} x, \qquad x \in \mathbb{R}^{n_1}, \lambda \in \mathbb{R}^{m_1}, \mu \in K^* \subseteq \mathbb{R}^{n_1}$$

Dual function:

$$d(\lambda, \mu) \;:=\; \min_x L(x, \lambda, \mu) \;=\; \lambda^\mathsf{T} b + d_0(\lambda, \mu)$$

$$d_0(\lambda, \mu) \;:=\; \min_{x \in \mathbb{R}^{n_1}} \left(f(x) + Q(x) - (A^\mathsf{T}\lambda + \mu)^\mathsf{T} x \right)$$

Dual problem:

$$d^* \;:=\; \max_{\lambda \in \mathbb{R}^{m_1},\, \mu \in K^*} \; \lambda^\mathsf{T} b + d_0(\lambda, \mu)$$

Strong duality and KKT

Nonsmooth conic program: $f^* \; := \; \min_{x \in \mathbb{R}^{n_1}} \; f(x) + Q(x) \quad \text{s.t.} \quad Ax = b,\; x \in K$

Assumptions

1. *Finite 2nd moment*: $E\zeta^2 < \infty$ and $Q(x) \in (-\infty, \infty]$
2. $f : \mathbb{R}^{n_1} \to \mathbb{R}$ is a convex function; K is a closed convex cone
3. *Slater condition*: $\exists \bar{x} \in \text{ri}(\text{dom}(Q)) \cap \text{ri}(K)$ such that $A\bar{x} = b$

Theorem [nonsmooth Slater theorem]

1. *Strong duality and dual optimality*: If f^* is finite, then $\exists$ dual optimal (λ^*, μ^*) that closes duality gap, i.e., $f^* = d^* = d(\lambda^*, \mu^*)$
2. *KKT characterization*: A feasible $x^* \in K$ with $Ax^* = b$ is primal optimal iff $\exists$ subgradients $\xi^* \in \partial f(x^*)$ and $\psi^* \in \partial Q(x^*)$, a dual feasible $(\lambda^*, \mu^*) \in \mathbb{R}^{m_1} \times K^*$ such that

$$\xi^* + \psi^* \; = \; A^{\mathsf{T}}\lambda^* + \mu^*, \qquad \mu^{*\mathsf{T}} x^* = 0$$

In this case (x^*, λ^*, μ^*) is a saddle point that closes the duality gap

Stochastic OPF

Summary

Brief introduction to theory of stochastic optimization

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad h(x, \zeta) \leq 0$$

where ζ is an uncertain parameter

Choose optimal x^* s.t.

- Robust opt: x^* satisfies constraints for all ζ in an uncertainty set Z
- Chance constrained opt: x^* satisfies constraints with high probability
- Scenario opt: x^* satisfies constraints for N random samples of $\zeta \in Z$
- Two-stage opt: 2nd-stage decision $y(x^*, \zeta)$ adapts to realized parameter ζ, given 1st-stage decision x^*

Many methods are combinations of these 4 ideas, e.g.

- Distributional robust opt: robust + chance constrained
- Adaptive robust opt: two-stage + robust (as opposed to expected) 2nd-stage cost
- Adaptive robust affine control: two-stage + robust (or avg) + affine policy