15.094J: Robust Modeling, Optimization, Computation

Lecture 6: Robust Convex Optimization

Outline

Motivation

- In earlier proposals (Ben-Tal and Nemirovski):

(a) RLOs become SOCPs

(b) Robust SOCPs become Semi-definite optimization problems (SDPs)

(c) Robust SDPs become NP-hard.

- In Lecture 4, we have shown that RLO becomes LO.
- In this lecture (Bertsimas and Sim), we show that :

$$\text{Robust SOCPs stay SOCPs}$$

$$\text{Robust SDPs stays SDPs}$$

- RC inherits the complexity of the underlying deterministic problem.
- RC allows the user to control the tradeoff between robustness and optimality.
- RC is computationally tractable both practically and theoretically.

Nominal vs Robust

- Nominal

$$\begin{aligned} \max \quad & f_0(\boldsymbol{x}, \tilde{\boldsymbol{D}}_0) \\ \text{s.t.} \quad & f_i(\boldsymbol{x}, \tilde{\boldsymbol{D}}_i) \geq 0, \qquad i \in I \\ & \boldsymbol{x} \in X \end{aligned}$$

- Exact Robust

$$\begin{aligned} \max \quad & \min_{\boldsymbol{D_0} \in \mathcal{U}_\mathbf{0}} f_0(\boldsymbol{x}, \boldsymbol{D}_0) \\ \text{s.t.} \quad & \min_{\boldsymbol{D}_i \in \mathcal{U}_i} f_i(\boldsymbol{x}, \boldsymbol{D}_i) \geq 0, i \in I \\ & \boldsymbol{x} \in X \end{aligned} \tag{1}$$

Uncertainty

- Data uncertainty

$$\tilde{\boldsymbol{D}} = \boldsymbol{D}^0 + \sum_{j \in N} \Delta \boldsymbol{D}^j \tilde{z}_j$$

- Uncertainty sets

$$\mathcal{U} = \left\{ \boldsymbol{D} \mid \exists \boldsymbol{u} \in \Re^{|N|} : \boldsymbol{D} = \boldsymbol{D}^0 + \sum_{j \in N} \Delta \boldsymbol{D}^j u_j, \|\boldsymbol{u}\| \leq \rho \right\}$$

Modeling power

Type	Constraint	D	$f(x, D)$
LO	$\boldsymbol{a}'\boldsymbol{x} \geq b$	$(\boldsymbol{a}, b)$	$\boldsymbol{a}'\boldsymbol{x} - b$
QCQO	$\lVert\boldsymbol{A}\boldsymbol{x}\rVert_2^2 + \boldsymbol{b}'\boldsymbol{x} + c \leq 0$	$(\boldsymbol{A}, \boldsymbol{b}, c, d)$ $d^0 = 1/2, \Delta d^j = 0$	$\frac{d-(\boldsymbol{b}'\boldsymbol{x}+c)}{2}$ $-\sqrt{\lVert\boldsymbol{A}\boldsymbol{x}\rVert_2^2 + \left(\frac{d+\boldsymbol{b}'\boldsymbol{x}+c}{2}\right)^2}$
SOCO(1)	$\lVert\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\rVert_2 \leq \boldsymbol{c}'\boldsymbol{x} + d$	$(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d)$ $\boldsymbol{\Delta c}^j = \boldsymbol{0}, \Delta d^j = 0$	$\boldsymbol{c}'\boldsymbol{x} + d - \lVert\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\rVert_2$
SOCO(2)	$\lVert\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\rVert_2 \leq \boldsymbol{c}'\boldsymbol{x} + d$	$(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d)$	$\boldsymbol{c}'\boldsymbol{x} + d - \lVert\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\rVert_2$
SDO	$\sum_{j=1}^n \boldsymbol{A}_i x_i - \boldsymbol{B} \in \boldsymbol{S}_+^m$	$(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n, \boldsymbol{B})$	$\lambda_{min}(\sum_{j=1}^n \boldsymbol{A}_i x_i - \boldsymbol{B})$

Exact and Relaxed Robustness

- Exact Robustness (ER)

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j \in N} \boldsymbol{\Delta D}^j u_j\right) \geq 0 \qquad \forall \|\boldsymbol{u}\| \leq \rho.$$

- Relaxed Robustness (RR)

$$f(\boldsymbol{x}, \boldsymbol{D}^0) + \sum_{j \in N} \left\{ f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) v_j + f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) w_j \right\} \geq 0$$

$$\forall (\boldsymbol{v}, \boldsymbol{w}) \in \Re_+^{|N| \times |N|} \ \|\boldsymbol{v} + \boldsymbol{w}\| \leq \rho.$$

Theorem

- Assumption 1: Norms satisfy $\|\boldsymbol{u}\| = \|\boldsymbol{u}^+\|$, $u_j^+ = |u_j|$. Examples L_p-norms.
- Assumption 2: f satisfies: $f(\boldsymbol{x}, \boldsymbol{D})$ is concave in $\boldsymbol{D}$ for all $\boldsymbol{x} \in \Re^n$, $f(\boldsymbol{x}, k\boldsymbol{D}) = kf(\boldsymbol{x}, \boldsymbol{D})$, for all $k \geq 0$, $\boldsymbol{D}$, $\boldsymbol{x} \in \Re^n$,

(a) Under Assumption 1 and $f(\boldsymbol{x}, \boldsymbol{A} + \boldsymbol{B}) = f(\boldsymbol{x}, \boldsymbol{A}) + f(\boldsymbol{x}, \boldsymbol{B})$, ER and RR are equivalent.

(b) Under Assumptions 1 and 2, if $\boldsymbol{x}^*$ satisfies RR, it satisfies ER also.

Proof of part (a)

- Under linearity, RR becomes

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j (v_j - w_j)\right) \geq 0 \qquad \forall \|\boldsymbol{v} + \boldsymbol{w}\| \leq \rho, \quad \boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0},$$

- ER becomes

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j r_j\right) \geq 0 \qquad \forall \|\boldsymbol{r}\| \leq \rho.$$

- If $\boldsymbol{x}$ violates ER, there exists $\boldsymbol{r}$, $\|\boldsymbol{r}\| \leq \rho$ such that

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j r_j\right) < 0.$$

- Let $v_j = \max\{r_j, 0\}$ and $w_j = -\min\{r_j, 0\}$.
- Clearly, $\boldsymbol{r} = \boldsymbol{v} - \boldsymbol{w}$ and since $v_j + w_j = |r_j|$, $\|\boldsymbol{v} + \boldsymbol{w}\| = \|\boldsymbol{r}\| \leq \rho$.
- $\boldsymbol{x}$ violates RR.

Proof of part (a), continued

- If $\boldsymbol{x}$ violates RR, then there exist $\boldsymbol{v}, \boldsymbol{w} \geq \mathbf{0}$ and $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \rho$ such that

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j \in N} \Delta \boldsymbol{D}^j (v_j - w_j)\right) < 0.$$

- Let $r_j = v_j - w_j$ and we observe that $|r_j| \leq v_j + w_j$.
- For norms satisfying $\|\boldsymbol{u}\| = \|\boldsymbol{u}^+\|$, $u_j^+ = |u_j|$,

$$\|\boldsymbol{r}\| = \|\boldsymbol{r}^+\| \leq \|\boldsymbol{v} + \boldsymbol{w}\| \leq \rho,$$

and hence, $\boldsymbol{x}$ is violates ER.

Proof of part (b)

- If $\boldsymbol{x}$ satisfies RR

$$f(\boldsymbol{x},\boldsymbol{D}^0)+\sum_{j\in N}\left\{f(\boldsymbol{x},\boldsymbol{\Delta D}^j)v_j+f(\boldsymbol{x},-\boldsymbol{\Delta D}^j)w_j\right\}\geq 0,\ \forall\|\boldsymbol{v}+\boldsymbol{w}\|\leq\rho,\ \ \boldsymbol{v},\boldsymbol{w}\geq\boldsymbol{0}.$$

- From concavity and homogeneity

$$f(\boldsymbol{x},\boldsymbol{A}+\boldsymbol{B})\geq\frac{1}{2}f(\boldsymbol{x},2\boldsymbol{A})+\frac{1}{2}f(\boldsymbol{x},2\boldsymbol{B})=f(\boldsymbol{x},\boldsymbol{A})+f(\boldsymbol{x},\boldsymbol{B}).$$

- Then

$$0\leq f(\boldsymbol{x},\boldsymbol{D}^0)+\sum_{j\in N}\left\{f(\boldsymbol{x},\boldsymbol{\Delta D}^j)v_j+f(\boldsymbol{x},-\boldsymbol{\Delta D}^j)w_j\right\}\leq$$

$$f(\boldsymbol{x},\boldsymbol{D}^0+\sum_{j\in N}\boldsymbol{\Delta D}^j(v_j-w_j))$$

for all $\|\boldsymbol{v}+\boldsymbol{w}\|\leq\rho,\quad\boldsymbol{v},\boldsymbol{w}\geq\boldsymbol{0}$.

Proof of part (b), continued

- In part (a) we established that

$$f(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j \in N} \boldsymbol{\Delta D}^j r_j) \geq 0 \qquad \forall \|\boldsymbol{r}\| \leq \rho$$

is equivalent to

$$f(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j \in N} \boldsymbol{\Delta D}^j (v_j - w_j)) \geq 0 \qquad \forall \|\boldsymbol{v} + \boldsymbol{w}\| \leq \rho, \quad \boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0},$$

and thus $\boldsymbol{x}$ satisfies ER.

Tractability

RR is equivalent to

$$\begin{aligned} & f(\boldsymbol{x}, \boldsymbol{D}^0) \geq \rho y \\ & f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) + t_j \geq 0 \qquad \forall j \in N \\ & f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) + t_j \geq 0 \quad \forall j \in N \\ & \|\boldsymbol{t}\|^* \leq y \\ & y \in \Re, \;\; \boldsymbol{t} \in \Re^{|N|}. \end{aligned}$$

Dual norm: $\|\boldsymbol{s}\|^* = \max_{\|\boldsymbol{x}\| \leq 1} \boldsymbol{s}'\boldsymbol{x}$.

Tractability, continued

(a) Under Assumption 1 and 2, RR is equivalent to RR'

$$f(\boldsymbol{x}, \boldsymbol{D}^0) \geq \rho \|\boldsymbol{s}\|^*,$$

where

$$s_j = \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)\}, \quad \forall j \in N.$$

(b) $f(\boldsymbol{x}, \boldsymbol{D}^0) \geq \rho \|\boldsymbol{s}\|^*$, can be written as:

$$\begin{aligned}
& f(\boldsymbol{x}, \boldsymbol{D}^0) \geq \rho y \\
& f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) + t_j \geq 0 \qquad \forall j \in N \\
& f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) + t_j \geq 0 \quad \forall j \in N \\
& \|\boldsymbol{t}\|^* \leq y \\
& y \in \Re, \ \boldsymbol{t} \in \Re^{|N|}.
\end{aligned}$$

Proof, part (a)

- We introduce the following problems:

$$\begin{aligned} z_1 = \max \quad & \boldsymbol{a}'\boldsymbol{v} + \boldsymbol{b}'\boldsymbol{w} \\ \text{s.t.} \quad & \|\boldsymbol{v} + \boldsymbol{w}\| \le \rho \\ & \boldsymbol{v}, \boldsymbol{w} \ge \boldsymbol{0}, \end{aligned}$$

$$\begin{aligned} z_2 = \max \quad & \sum_{j \in N} \max\{a_j, b_j, 0\} r_j \\ \text{s.t.} \quad & \|\boldsymbol{r}\| \le \rho, \end{aligned}$$

and show that $z_1 = z_2$.

- Suppose $\boldsymbol{r}^*$ is an optimal solution to z_2. For all $j \in N$, let

$$\begin{aligned} v_j = w_j = 0 \qquad & \text{if } \max\{a_j, b_j\} \le 0 \\ v_j = |r_j^*|, w_j = 0 \quad & \text{if } a_j \ge b_j, a_j > 0 \\ w_j = |r_j^*|, v_j = 0 \quad & \text{if } b_j > a_j, b_j > 0. \end{aligned}$$

Proof part (a), continued

- Observe that $a_j v_j + b_j w_j \geq \max\{a_j, b_j, 0\} r_j^*$ and $w_j + v_j \leq |r_j^*|\ \forall j \in N$.
- If $\boldsymbol{v}^+ \leq \boldsymbol{w}^+$, $\|\boldsymbol{v}\| \leq \|\boldsymbol{w}\|$.
- Then $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{r}^*\| \leq \rho$, and thus $\boldsymbol{v}, \boldsymbol{w}$ are feasible in z_1 leading to

$$z_1 \geq \sum_{j \in N} (a_j v_j + b_j w_j) \geq \sum_{j \in N} \max\{a_j, b_j, 0\} r_j^* = z_2.$$

- Conversely, let $\boldsymbol{v}^*, \boldsymbol{w}^*$ be an optimal solution to z_1.
- Let $\boldsymbol{r} = \boldsymbol{v}^* + \boldsymbol{w}^*$. Clearly $\|\boldsymbol{r}\| \leq \rho$ and observe that

$$r_j \max\{a_j, b_j, 0\} \geq a_j v_j^* + b_j w_j^* \quad \forall j \in N.$$

- Therefore, we have

$$z_2 \geq \sum_{j \in N} \max\{a_j, b_j, 0\} r_j \geq \sum_{j \in N} (a_j v_j^* + b_j w_j^*) = z_1,$$

leading to $z_1 = z_2$.

Proof part (a), continued

- $\mathcal{V} = \{(\boldsymbol{v}, \boldsymbol{w}) \in \Re_+^{|N| \times |N|} \; \|\boldsymbol{v} + \boldsymbol{w}\| \leq \rho\}$.
- Then,

$$\begin{aligned}
& \min_{(\boldsymbol{v},\boldsymbol{w}) \in \mathcal{V}} \sum_{j \in N} \left\{ f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) v_j + f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) w_j \right\} \\
= \quad & - \max_{(\boldsymbol{v},\boldsymbol{w}) \in \mathcal{V}} \sum_{j \in N} \left\{ -f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) v_j - f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) w_j \right\} \\
= \quad & - \max_{\{\|\boldsymbol{r}\| \leq \rho\}} \sum_{j \in N} \left\{ \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j), 0\} r_j \right\}
\end{aligned}$$

- Since $\|\boldsymbol{s}\|^* = \max_{\|x\| \leq 1} \boldsymbol{s}'\boldsymbol{x}$, we obtain $\rho\|\boldsymbol{s}\|^* = \max_{\|x\| \leq \rho} \boldsymbol{s}'\boldsymbol{x}$, i.e., RR' follows.
- Note that $s_j = \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)\} \geq 0$, since otherwise there exists an $\boldsymbol{x}$ such that $s_j < 0$, i.e., $f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) > 0$ and $f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) > 0$. From Assumption 2 $f(\boldsymbol{x}, \boldsymbol{0}) = 0$, contradicting the concavity of $f(\boldsymbol{x}, \boldsymbol{D})$.

Proof, part (b)

- Suppose that $\boldsymbol{x}$ is feasible in RR'.
- Let $\boldsymbol{t} = \boldsymbol{s}$ and $y = \|\boldsymbol{s}\|^*$,
- We can easily check that $(\boldsymbol{x}, \boldsymbol{t}, y)$ are feasible in RR".
- Conversely, suppose, $\boldsymbol{x}$ is infeasible in RR', that is,

$$f(\boldsymbol{x}, \boldsymbol{D}^0) < \rho\|\boldsymbol{s}\|^*.$$

- Since, $t_j \geq s_j = \max\{-f(\boldsymbol{x}, \Delta\boldsymbol{D}^j), -f(\boldsymbol{x}, -\Delta\boldsymbol{D}^j)\} \geq 0$
- We have $\boldsymbol{v}^+ \leq \boldsymbol{w}^+, \ \|\boldsymbol{v}\|^* \leq \|\boldsymbol{w}\|^*$.
- Thus, $\|\boldsymbol{t}\|^* \geq \|\boldsymbol{s}\|^*$, leading to

$$f(\boldsymbol{x}, \boldsymbol{D}^0) < \rho\|\boldsymbol{s}\|^* \leq \rho\|\boldsymbol{t}\|^* \leq \rho y,$$

  i.e., $\boldsymbol{x}$ is infeasible in RR".

Dual norm

Norms	$\|\boldsymbol{u}\|$	$\|\boldsymbol{t}\|^* \leq y$
L_2	$\|\boldsymbol{u}\|_2$	$\|\boldsymbol{t}\|_2 \leq y$
L_1	$\|\boldsymbol{u}\|_1$	$t_j \leq y, \forall j \in N$
L_∞	$\|\boldsymbol{u}\|_\infty$	$\sum_{j \in N} t_j \leq y$
L_p	$\|\boldsymbol{u}\|_p$	$\left(\sum_{j \in N} t_j^{\frac{q}{q-1}}\right)^{\frac{q-1}{q}} \leq y$
$L_2 \cap L_\infty$	$\max\{\|\boldsymbol{u}\|_2, \rho\|\boldsymbol{u}\|_\infty\}$	$\|\boldsymbol{s} - \boldsymbol{t}\|_2 +$ $\frac{1}{\rho}\sum_{j \in N} s_j \leq y,\ \ \boldsymbol{s} \in \Re_+^{\lvert N \rvert}$
$L_1 \cap L_\infty$	$\max\{\frac{1}{\Gamma}\|\boldsymbol{u}\|_1, \|\boldsymbol{u}\|_\infty\}$	$\Gamma p + \sum_{j \in N} s_j \leq y$ $s_j + p \geq t_j,\ \ p \in \Re_+, \boldsymbol{s} \in \Re_+^{\lvert N \rvert}$

Size

- Independent Pertubations
- Example

$$\begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} x_1 + \begin{pmatrix} a_4 & a_5 \\ a_5 & a_6 \end{pmatrix} x_2 \succeq \begin{pmatrix} a_7 & a_8 \\ a_8 & a_9 \end{pmatrix},$$

$\tilde{a}_i = a_i^0 + \Delta a_i \tilde{z}_i$.

- $f(\boldsymbol{x}, \boldsymbol{\Delta d}^1) + t_1 \geq 0$ becomes

$$\lambda_{min}\left(\begin{pmatrix} \Delta a_1 & 0 \\ 0 & 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} x_2 - \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right) + t_1 \geq 0,$$

as $t_1 \geq -\min\{\Delta a_1 x_1, 0\}$ or equivalently as linear constraints $t_1 \geq -\Delta a_1 x_1, t_1 \geq 0$.

Tractability

	L_∞	L_1	L_2	$L_2 \cap L_\infty$
Num. Vars.	$n+1$	1	1	$2\|N\|+1$
Num. linear Const.	$2n+1$	$2n+1$	0	$3\|N\|$
Num SOC Const.	0	0	1	1
LO	LO	LO	SOCO	SOCO
QCQO	SOCO	SOCO	SOCO	SOCO
SOCO(1)	SOCO	SOCO	SOCO	SOCO
SOCO(2)	SOCO	SOCO	SOCO	SOCO
SDO	SDO	SDO	SDO	SDO

Probabilistic Guarantees

If $\tilde{\boldsymbol{z}} \sim \mathcal{N}(0, \boldsymbol{I})$, under the L_2 norm:

$$\mathrm{P}(f(\boldsymbol{x}, \tilde{\boldsymbol{D}}) < 0) \leq \frac{\sqrt{e}\rho}{\alpha} e^{\left(-\frac{\rho^2}{2\alpha^2}\right)}$$

Problem	α	ρ
LO	1	$O(log(1/\epsilon))$
SOCO(1)	1	$O(log(1/\epsilon))$
SOCO(2)	$\sqrt{2}$	$O(log(1/\epsilon))$
QCQO	$\sqrt{2}$	$O(log(1/\epsilon))$
SDO	$\sqrt{\log m}$	$O(\sqrt{\log m} log(1/\epsilon))$

Conclusions

- Given a conic optimization problem, we proposed a robust counterpart of the same character as original, thus preserving computational tractability.
- Size of the proposed problem is very similar to original; depends on the norm we use; best results for L_2 norm.
- Probabilistic guarantee allows to select parameter controlling robustness and optimality.