15.094J: Robust Modeling, Optimization, Computation

Lecture 6: Robust Convex Optimization

February 2015

Outline

1. Motivation
2. Robust Conic Optimization
3. Exact and Relaxed Robustness
4. Tractability
5. Probabilistic Guarantees
6. Conclusions

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 Contrast
  - (a) In Lecture 4, we have shown that RLO becomes LO.
  - (b) Today we show that Robust SOCPs stay SOCPs
  - (c) and Robust SDPs stay 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

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

- Exact Robust

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

Uncertainty

- Data uncertainty

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

- Uncertainty sets

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

Modeling power

Type	Constraint	D	$f(x, D)$
LO	$a'x \geq b$	(a, b)	$a'x - b$
QCQO	$\|Ax\|_2^2 + b'x + c \leq 0$	(A, b, c, d) $d^0 = 1/2, \Delta d^j = 0$	$\frac{d-(b'x+c)}{2}$ $-\sqrt{\|Ax\|_2^2 + \left(\frac{d+b'x+c}{2}\right)^2}$
SOCO(1)	$\|Ax + b\|_2 \leq c'x + d$	(A, b, c, d) $\Delta c^j = \mathbf{0}, \Delta d^j = 0$	$c'x + d - \|Ax + b\|_2$
SOCO(2)	$\|Ax + b\|_2 \leq c'x + d$	(A, b, c, d)	$c'x + d - \|Ax + b\|_2$
SDO	$\sum_{j=1}^n A_i x_i - B \in S_+^m$	$(A_1, ..., A_n, B)$	$\lambda_{min}(\sum_{j=1}^n A_i x_i - B)$

Exact and Relaxed Robustness

- Exact Robustness (ER)

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

- Relaxed Robustness (RR)

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

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

Theorem

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

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

(b) Under Assumptions 1 and 2, if x^* satisfies RR, it satisfies ER also.

Proof of part (a)

- Under linearity, RR becomes

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

- ER becomes

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

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

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

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

Proof of part (a), continued

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

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

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

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

and hence, x violates ER.

Proof of part (b)

- If x satisfies RR

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

- From concavity and homogeneity

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

- Then

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

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

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

Proof of part (b), continued

- In part (a) we established that

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

is equivalent to

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

and thus x satisfies ER.

Tractability

RR is equivalent to

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

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

Tractability, continued

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

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

where

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

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

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

Proof, part (a)

- We introduce the following problems:

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

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

  and show that $z_1 = z_2$.

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

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

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 $v^+ \leq w^+$, $\|v\| \leq \|w\|$.
- Then $\|v + w\| \leq \|r^*\| \leq \rho$, and thus v, 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 v^*, w^* be an optimal solution to z_1.
- Let $r = v^* + w^*$. Clearly $\|r\| \leq \rho$ 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

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

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

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

Proof, part (b)

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

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

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

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

i.e., x is infeasible in RR".

Dual norm

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

Size

- Independent Pertubations
- Example

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

  $\tilde{a}_i = a_i^0 + \Delta a_i \tilde{z}_i$.
- $f(x, \boldsymbol{\Delta} d^1) + t_1 \geq 0$ becomes

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

  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	$2\lvert N\rvert+1$
Num. linear Const.	$2n+1$	$2n+1$	0	$3\lvert N\rvert$
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 $\tilde{z} \sim \mathcal{N}(0, I)$, under the L_2 norm:

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

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.