Linear & Conic Programming Reformulations of Two-Stage Robust Linear Programs

Erick Delage
CRC in decision making under uncertainty
Department of Decision Sciences
HEC Montreal

(joint work with Amir Ardestani-Jaafari)
(special thanks to Samuel Burer)

July 5th, 2017
A Classical Distribution Problem

- A multinational retailing corporation wishes to construct new warehouses
A Classical Distribution Problem

1. Choose where to build the new warehouses
A CLASSICAL DISTRIBUTION PROBLEM

2. Observe amount of weekly demand

![Map with warehouse sites and retailers](image-url)
**A Classical Distribution Problem**

3. Transport goods to retailers to maximize profits
A Classical Distribution Problem

- Facility location-transportation model

\[
\begin{align*}
\text{maximize} & \quad \eta \sum_{i} \sum_{j} Y_{ij} - \left( c^T x + K^T I + \sum_{i} \sum_{j} (p_i + t_{ij}) Y_{ij} \right) \\
\text{s.t.} & \quad \sum_{j} Y_{ij} \leq x_i, \ \forall \ i, \quad \text{(Capacity constraint)} \\
& \quad \sum_{i} Y_{ij} \leq d_j, \ \forall \ j, \quad \text{(Demand constraint)} \\
& \quad x_i \leq M I_i, \ \forall \ i, \quad \text{(Facility Size constraint)}
\end{align*}
\]

How can one account for demand uncertainty?
Robust Optimization is Now a Well Established Methodology
A Classical Robust Distribution Problem

- Robust Facility location-transportation model:

\[
\begin{align*}
\text{maximize} & \quad \min_{I \in \{0,1\}^n, x} h(I, x, d) \\
\text{s. t.} & \quad x_i \leq MI_i, \; \forall \; i, \quad (\text{Facility Size constraint})
\end{align*}
\]

where \( h(I, x, d) \) is the optimal value of

\[
\begin{align*}
\max_{Y \geq 0} & \quad \eta \sum_i \sum_j Y_{ij} - \left( c^T x + K^T I + \sum_i \sum_j (p_i + t_{ij}) Y_{ij} \right) \\
\text{s. t.} & \quad \sum_j Y_{ij} \leq x_i, \; \forall \; i, \quad (\text{Capacity constraint}) \\
& \quad \sum_i Y_{ij} \leq d_j, \; \forall \; j, \quad (\text{Demand constraint})
\end{align*}
\]
OUTLINE

INTRODUCTION

BACKGROUND ON TWO-STAGE ROBUST LINEAR PROGRAMS

COPOSITIVE PROGRAMMING REFORMULATIONS

APPLICATIONS

CONCLUSION
OUTLINE

INTRODUCTION

BACKGROUND ON TWO-STAGE ROBUST LINEAR PROGRAMS

COPOSITIVE PROGRAMMING REFORMULATIONS

APPLICATIONS

CONCLUSION
**Static Robust Linear Program**

[Ben-Tal & Nemirovski (2000), 1296 citations !]

- Consider the following static problem:

  \[
  \text{maximize} \quad c^T x + f^T y \\
  \text{s. t.} \quad Ax + By \leq D(x)z, \forall z \in \mathcal{Z} 
  \]

  where we assume \( n_x + n_y \) decision variables, \( J \) constraints, and \( m \) uncertain parameters.

- If \( \mathcal{Z} := \{z \in \mathbb{R}^m \mid z \geq 0, Pz = q\} \) is a non-empty polyhedral set defined by \( K \) constraints, then

  \[
  \text{Problem (1)} \equiv \text{maximize} \quad c^T x + f^T y \\
  \text{s. t.} \quad Ax + By + \Lambda q \leq 0 \\
  \quad D(x) + \Lambda P \geq 0,
  \]

  where \( \Lambda \in \mathbb{R}^{J \times K} \).
TWO-STAGE ROBUST LINEAR PROGRAMS

[Ben-Tal et al. (2004), 824 citations!]

- Consider the following two-stage problem:

\[
\begin{align*}
\text{(TSRLP)} & \quad \text{maximize} \quad x \in X, y(\cdot) \quad \min_{z \in Z} \ c^T x + f^T y(z) \\
& \quad \text{s. t.} \quad Ax + By(z) \leq D(x)z \ \forall z \in Z
\end{align*}
\]

where \( y : \mathbb{R}^m \rightarrow \mathbb{R}^{ny} \)

- This problem can also be represented as

\[
\begin{align*}
\text{(TSRLP)} & \quad \text{maximize} \quad x \in X \quad \min_{z \in Z} \ h(x, z) \\
& \quad \text{where} \quad h(x, z) := \max_y \ c^T x + f^T y \\
& \quad \text{s. t.} \quad Ax + By \leq D(x)z.
\end{align*}
\]
**Complexity of Two-Stage Robust Linear Programs**

- Unfortunately, the two-stage robust linear program is known to be intractable in general [Ben-Tal et al. (2004)].

- Conservative approximation obtained by using affine adjustment functions:

  \[ y(z) := y + Yz \]

  The two-stage robust problem reduces to

  \[
  (AARC) \quad \max_{x \in \mathcal{X}, y, Y} \min_{z \in \mathcal{Z}} \quad c^T x + f^T (y + Yz)
  \]

  s. t. \quad Ax + B(y + Yz) \leq D(x)z \quad \forall z \in \mathcal{Z}

- Some exact methods have been proposed but without polynomial time convergence guarantees [Zeng & Zhao (2013)]
OUTLINE

INTRODUCTION

BACKGROUND ON TWO-STAGE ROBUST LINEAR PROGRAMS

COPOSITIVE PROGRAMMING REFORMULATIONS

APPLICATIONS

CONCLUSION
**COPOSITIVE PROGRAMMING REFORMULATION I**

- **Assumptions**
  1. $\mathcal{Z}$ is a non-empty and bounded polyhedral set
  2. The TSRLP problem is bounded above, i.e.
     \[ \forall x \in \mathcal{X}, \exists z \in \mathcal{Z}, h(x, z) < \infty. \]

- **Let our robust optimization problem take the form**
  \[
  \max_{x \in \mathcal{X}} \psi(x),
  \]
  where
  \[
  \psi(x) := \min_{z \in \mathcal{Z}} \max_{y} c^T x + f^T y \tag{2a}
  \]
  s. t. \[ Ax + By \leq D(x)z \tag{2b} \]

- **Since (2) is bounded, strong LP duality applies**
  \[
  \psi(x) = \min_{z \in \mathcal{Z}, \lambda \geq 0} c^T x + z^T D(x)^T \lambda - (Ax)^T \lambda
  \]
  \[ B^T \lambda = f \]
The function $\psi(x)$ minimizes a non-convex quadratic function over a polyhedron in the non-negative orthant

$$
\psi(x) = \min_{\tilde{z} \geq 0} \quad c^T x + \tilde{z}^T \tilde{Q}(x)\tilde{z} - \tilde{c}(x)^T \tilde{z}
$$

$$
\tilde{A}\tilde{z} = \tilde{b},
$$

where $\tilde{z} := [\lambda^T \quad z^T] \in \mathbb{R}^{J+m}$ and where

$$
\tilde{Q}(x) := \begin{bmatrix}
0 & (1/2)D(x) \\
(1/2)D(x)^T & 0
\end{bmatrix} \quad \tilde{c}(x) := \begin{bmatrix}
-(1/2)A x \\
0
\end{bmatrix}
$$

$$
\tilde{A} := \begin{bmatrix}
B^T & 0 \\
0 & P
\end{bmatrix} \quad \tilde{b} := \begin{bmatrix}
d \\
q
\end{bmatrix}
$$
**Copositive Programming Reformulation I**

- The function $\psi(x)$ minimizes a non-convex quadratic function over a polyhedron in the non-negative orthant

$$\psi(x) = \min_{\tilde{z} \geq 0} \ c^T x + trace(\tilde{Q}(x)^T \tilde{z} \tilde{z}^T) - \tilde{c}(x)^T \tilde{z}$$

$$\tilde{A} \tilde{z} = \tilde{b}$$

$$\tilde{A} \tilde{z} \tilde{z}^T = \tilde{b} \tilde{z}^T,$$

where $\tilde{z} := [\lambda^T \ z^T] \in \mathbb{R}^{J+m}$ and where

$$\tilde{Q}(x) := \begin{bmatrix} 0 & (1/2)D(x) \\ (1/2)D(x)^T & 0 \end{bmatrix}, \quad \tilde{\mathbf{c}}(x) := \begin{bmatrix} -(1/2)A\mathbf{x} \\ 0 \end{bmatrix}$$

$$\tilde{\mathbf{A}} := \begin{bmatrix} B^T & 0 \\ 0 & P \end{bmatrix}, \quad \tilde{b} := \begin{bmatrix} d \\ q \end{bmatrix}$$
The function $\psi(x)$ has an equivalent convex optimization reformulation ($\tilde{Z} := \tilde{z}\tilde{z}^T$) [Burer (2009)]

$$\psi(x) = \min_{\tilde{Z}, \tilde{z}} c^T x + \text{trace}(\tilde{Q}(x)^T\tilde{Z}) - \tilde{c}(x)^T\tilde{z}$$

$$\tilde{A}\tilde{z} = \tilde{b}$$

$$\tilde{A}\tilde{Z} = \tilde{b}\tilde{z}^T$$

$$\begin{bmatrix} \tilde{Z} & \tilde{z} \\ \tilde{z}^T & 1 \end{bmatrix} \in \mathcal{K}_{\text{CP}} \quad \text{and} \quad \text{rank} \left( \begin{bmatrix} \tilde{Z} & \tilde{z} \\ \tilde{z}^T & 1 \end{bmatrix} \right) = 1$$

where $\mathcal{K}_{\text{CP}}$ is the cone of completely positive matrices, i.e.

$$\mathcal{K}_{\text{CP}} := \left\{ M \left| M = \sum_{k \in K} \tilde{z}_k\tilde{z}_k^T \text{ for some } \{\tilde{z}_k\}_{k \in K} \subset \mathbb{R}_+^{J+m+1} \right. \right\}.$$
The function $\psi(x)$ has an equivalent convex optimization reformulation ($\tilde{Z} := \tilde{z}\tilde{z}^T$) [Burer (2009)]

$$\psi(x) = \min_{\tilde{Z}, \tilde{z}} c^T x + \text{trace}(\tilde{Q}(x)^T\tilde{Z}) - \tilde{c}(x)^T\tilde{z}$$

$$\tilde{A}\tilde{z} = \tilde{b}$$

$$\tilde{A}\tilde{Z} = \tilde{b}\tilde{z}^T$$

$$\begin{bmatrix} \tilde{Z} & \tilde{z} \\ \tilde{z}^T & 1 \end{bmatrix} \in \mathcal{K}_{\text{CP}}$$

where $\mathcal{K}_{\text{CP}}$ is the cone of completely positive matrices, i.e.

$$\mathcal{K}_{\text{CP}} := \left\{ M \bigg| M = \sum_{k \in K} \tilde{z}_k\tilde{z}_k^T \text{ for some } \{\tilde{z}_k\}_{k \in K} \subset \mathbb{R}_+^{J+m+1} \right\}.$$
By conic duality we get

\[ \psi(x) \geq \max_{\tilde{W}, \tilde{v}, \tilde{w}, t} \tilde{c}(x)^T x + \tilde{b}^T \tilde{w} - t \]

\[ \text{s. t. } \tilde{v} = \tilde{c}(x) - (1/2)(\tilde{A}^T \tilde{w} - \tilde{W}^T \tilde{b}) \]

\[ \begin{bmatrix} \tilde{Q}(x) - (1/2)(\tilde{W}^T \tilde{A} + \tilde{A}^T \tilde{W}) & \tilde{v} \\ \tilde{v}^T & t \end{bmatrix} \in \mathcal{K}_{\text{Cop}}, \]

where \( \mathcal{K}_{\text{Cop}} \) is the cone of copositive matrices, i.e.

\[ \mathcal{K}_{\text{Cop}} := \left\{ M \bigg| M = M^T, \ z^T M z \geq 0, \ \forall z \in \mathbb{R}_+^{J+m+1} \right\}. \]
**COPOSITIVE PROGRAMMING REFORMULATION I**

**Theorem 1**  
[Xu & Burer (2016), Hanasusanto & Kuhn (2016)]  
If the TSRLP problem has “complete recourse”, i.e.

\[ \exists y \in \mathbb{R}^{n_y}, By < 0, \]

then the copositive program

\[
\begin{align*}
\text{(Copos}_1) \quad & \text{maximize} \quad c^T x + \tilde{b}^T \tilde{v} - t \\
& \text{s.t.} \quad \tilde{v} = \tilde{c}(x) - (1/2)(\tilde{A}^T \tilde{w} - \tilde{W}^T \tilde{b}) \\
& \left[ \begin{array}{c}
\tilde{Q}(x) - (1/2)(\tilde{W}^T \tilde{A} + \tilde{A}^T \tilde{W}) \\
\tilde{v}^T \\
t
\end{array} \right] \in \mathcal{K}_{Cop},
\end{align*}
\]

provides an exact reformulation of the TSRLP problem. Otherwise, Copos$_1$ only provides a conservative approximation.
Relation to AARC

Theorem 2  [Xu & Burer (2016)]
When $\mathcal{K}_{\text{Cop}}$ is replaced with $\mathcal{N} := \mathbb{R}_+^{J+m+1 \times J+m+1} \subset \mathcal{K}_{\text{Cop}}$ the copositive programming reformulation is equivalent to AARC.

- Hence, for any cone $\mathcal{K}$ such that $\mathcal{N} \subset \mathcal{K} \subset \mathcal{K}_{\text{Cop}}$, Copos$_1$ with $\mathcal{K}$ provides a tighter approximation than AARC.
- There exists a hierarchy of semidefinite and polyhedral cones $\{\mathcal{K}_i\}_{i=1}^{\infty}$, with $\mathcal{N} \subseteq \mathcal{K}_1 \subset \mathcal{K}_2 \subset \cdots \subset \mathcal{K}_{\text{Cop}}$, such that for all $M \in \mathcal{K}_{\text{Cop}}$, there is a $i^*, M \in \mathcal{K}_{i^*}$ [Parrilo (2000), Bomze & de Klerk (2002)]

- This is valuable for complete recourse problems but what about relatively complete recourse problems?
**HOW TO FIX RELATIVELY COMPLETE recourse**

- **Assumption**: The TSRLP problem has relatively complete recourse, i.e.

\[ \forall x \in \mathcal{X} \forall z \in \mathcal{Z}, \exists y, Ax + By \leq D(x)z \]

- This ensures that:

\[
h(x, z) = \min_{\lambda} \quad c^T x + z^T D(x)^T \lambda - (Ax)^T \lambda \quad \in \mathbb{R}
\]

\[
s. t. \quad \lambda \in \mathcal{P} := \{\lambda | \lambda \geq 0, B^T \lambda = f\}
\]

- Hence, always an optimal solution \( \lambda^*(x, z) \) at a vertex of \( \mathcal{P} \)

- Since number of vertices is finite, there exists \( u \in \mathbb{R}^J_+ \):

\[
\psi(x) = \min_{z \in \mathcal{Z}} h(x, z) = \min_{z \in \mathcal{Z}, \lambda \in \mathcal{P}} \quad c^T x + z^T D(x)^T \lambda - (Ax)^T \lambda
\]

\[
s. t. \quad \lambda \leq u
\]
The function $\psi(x)$ minimizes a non-convex quadratic function over a polyhedron in the non-negative orthant

$$
\psi(x) = \min_{\bar{y} \geq 0} \quad c^T x + \bar{y}^T \bar{Q}(x) \bar{y} - \bar{c}(x)^T \bar{y}
$$

$$
\bar{A} \bar{y} = \bar{b},
$$

where $\bar{y} := [\lambda^T \quad z^T \quad s^T] \in \mathbb{R}^{2J+m}$ and where

$$
\bar{Q}(x) := \begin{bmatrix}
0 & (1/2)D(x) & 0 \\
(1/2)D(x)^T & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
$$

$$
\bar{c}(x) := \begin{bmatrix}
-(1/2)Ax \\
0 \\
0
\end{bmatrix}
$$

$$
\bar{A} := \begin{bmatrix}
B^T & 0 & 0 \\
0 & P & 0 \\
I & 0 & I
\end{bmatrix}
$$

$$
\bar{b} := \begin{bmatrix}
d \\
q \\
u
\end{bmatrix}.
$$
Copositive Programming Reformulation II

Theorem 3  [AJ&D (2016b)]
If the TSRLP problem has relatively complete recourse, then the copositive program

$$(\text{Copos}_2) \quad \max_{x \in X, \bar{W}, \bar{w}, \bar{v}, t} \quad c^T x + \bar{b}^T \bar{w} - t$$

s. t.  \quad \bar{v} = \bar{c}(x) - (1/2)(\bar{A}^T \bar{w} - \bar{W}^T \bar{b})

$$\begin{bmatrix} \bar{Q}(x) - (1/2)(\bar{W}^T \bar{A} + \bar{A}^T \bar{W}) \\ \bar{v}^T \\ t \end{bmatrix} \in K_{\text{Cop}},$$

provides an exact reformulation of the TSRLP problem.
THE PENALIZED AARC MODEL

Theorem 4  [AJ&D (2016b)]
When $K_{Cop}$ is replaced with $N$ the Copos$_2$ reformulation is equivalent to applying affine adjustments to:

$$(TSRLP') \begin{align*}
& \text{maximize} \quad x^T x + f^T y(z) - u^T \theta(z) \\
& \text{min} \quad \min_{z \in \mathcal{Z}} \\
& \text{s. t.} \quad Ax + By(z) \leq D(x)z + \theta(z) \quad \forall z \in \mathcal{Z}.
\end{align*}$$

Moreover, affine (and static) adjustments are always feasible in TSRLP'.

- $u$ can be interpreted as a marginal penalty for violating constraints
- TSRLP' $\equiv$ TSRLP since $u$ is such that there always exists an optimal solution triplet with $\theta(z) := 0$.
- Method for converting a relatively complete recourse multi-stage linear program into a complete recourse one
OUTLINE

INTRODUCTION

BACKGROUND ON TWO-STAGE ROBUST LINEAR PROGRAMS

COPOSITIVE PROGRAMMING REFORMULATIONS

APPLICATIONS

CONCLUSION
**Robust Facility Location - Transportation Problem**

- In AJ&D (2016b), we identify an instance for which

<table>
<thead>
<tr>
<th></th>
<th>AARC model</th>
<th>Penalized AARC (a.k.a. Copos₂(ℕ))</th>
<th>Exact model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bound on w.-c. profit</strong></td>
<td>0</td>
<td>6600</td>
<td>6600</td>
</tr>
<tr>
<td><strong>W.-c. profit of x</strong></td>
<td>0</td>
<td>6600</td>
<td>6600</td>
</tr>
</tbody>
</table>

- We recently randomly generated 10 000 problem instances, 5 facilities & 10 customer locations.

<table>
<thead>
<tr>
<th>Optimality gap</th>
<th>Proportion of instances</th>
<th>Proportion of instances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AARC</td>
<td>Penalized AARC</td>
</tr>
<tr>
<td><strong>= 0%</strong></td>
<td>20.6%</td>
<td>23.8%</td>
</tr>
<tr>
<td><strong>≤ 0.1%</strong></td>
<td>20.9%</td>
<td>27.4%</td>
</tr>
<tr>
<td><strong>≤ 1%</strong></td>
<td>28.4%</td>
<td>56.3%</td>
</tr>
<tr>
<td><strong>Avg. Gap</strong></td>
<td>10.5%</td>
<td>1.6%</td>
</tr>
<tr>
<td><strong>Max Gap</strong></td>
<td>50.0%</td>
<td>13.3%</td>
</tr>
</tbody>
</table>
## WHAT SIZE PROBLEMS CAN WE SOLVE? [AJ&D (2017)]

<table>
<thead>
<tr>
<th>(T,L,N)</th>
<th>Γ</th>
<th>Penalized AARC</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Full form</td>
<td>Row generation</td>
</tr>
<tr>
<td>(1,50,100)</td>
<td>10</td>
<td>-</td>
<td>3 241 sec</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>-</td>
<td>4 563 sec</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>-</td>
<td>8 460 sec</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>-</td>
<td>3 781 sec</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>-</td>
<td>1 382 sec</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>-</td>
<td>&lt; 1 sec</td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>-</td>
<td>3 572 sec</td>
</tr>
<tr>
<td>(20,15,30)</td>
<td>60</td>
<td>-</td>
<td>3 781 sec</td>
</tr>
<tr>
<td></td>
<td>180</td>
<td>-</td>
<td>5 646 sec</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>-</td>
<td>10 567 sec</td>
</tr>
<tr>
<td></td>
<td>420</td>
<td>-</td>
<td>4 445 sec</td>
</tr>
<tr>
<td></td>
<td>540</td>
<td>-</td>
<td>663 sec</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>-</td>
<td>1 sec</td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>-</td>
<td>4 184 sec</td>
</tr>
</tbody>
</table>

(− stands for more than two days of computation)
ROBUST MULTI-ITEM NEWSVENDOR

- In AJ&D (2016a): the robust multi-item newsvendor problem with uncorrelated demand can be solved optimally by AARC/Copos(\(\mathcal{N}\)) when using budgeted uncertainty set with integer \(\Gamma\).

- In AJ&D (2016b): if demand is correlated than solution improves using Copos with \(\mathcal{K} \supset \mathcal{N}\):

<table>
<thead>
<tr>
<th></th>
<th>AARC</th>
<th>(\text{Copos}(\mathcal{K}_{\text{LP}}^4))</th>
<th>(\text{Copos}(\mathcal{K}_{\text{SDP}}^1))</th>
<th>Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>W.-c. profit bound</td>
<td>41.83</td>
<td>41.83</td>
<td>411.08</td>
<td>825.83</td>
</tr>
<tr>
<td>Actual w.-c. profit</td>
<td>41.83</td>
<td>41.83</td>
<td>664.76</td>
<td>825.83</td>
</tr>
</tbody>
</table>
OUTLINE

INTRODUCTION

BACKGROUND ON TWO-STAGE ROBUST LINEAR PROGRAMS

COPOSITIVE PROGRAMMING REFORMULATIONS

APPLICATIONS

CONCLUSION
CONCLUSION & OPEN QUESTIONS

1. Copositive programming is a useful tool for generating conservative approximations for TSRLP
   - $\text{Copos}(\mathcal{K})$ with $\mathcal{K} \supset \mathcal{N}$ always improves on AARC
   - Although hierarchy of polyhedral cones $\mathcal{N} \subset \mathcal{K}_{LP}^d \subset \mathcal{K}_{\text{Cop}}$ provide LP reformulations, preliminary results indicate that classical ones perform poorly
     - Can $\text{Copos}(\mathcal{K})$ provide intuition on approximate policies?
     - Do $\text{Copos}(\mathcal{K})$ reformulations exist for multi-stage problems?

2. Penalized violations transform any two-stage LP with relatively complete recourse in one with complete recourse
   - A useful preprocessing step for AARC when feasibility is a challenge
     - Is it possible to generalize this approach to robust multi-stage non-linear problems?
BIBLIOGRAPHY


► A. Ardestani-Jaafari and E. Delage. Linearized Robust Counterparts of Two-stage Robust Optimization Problem with Applications in Operations Management. Working draft, 2016b.

► A. Ardestani-Jaafari and E. Delage. The Value of Flexibility in Robust LocationTransportation Problems. Transportation Science, 2017.


Questions & Comments ...

... Thank you!