Linear Optimization Problem, its Data and Structure

♣ Linear Optimization problem:

\[
\min_x \{ c^T x + d : A x \leq b \} \tag{LO}
\]

- \( x \in \mathbb{R}^n \): vector of decision variables,
- \( c \in \mathbb{R}^n \) and \( d \in \mathbb{R} \) form the objective,
- \( A \): an \( m \times n \) constraint matrix,
- \( b \in \mathbb{R}^m \): right hand side.

♠ Problem’s structure: its sizes \( m, n \).
♠ Problem’s data: \((c, d, A, b)\).

Data Uncertainty

♣ The data of typical real world LOs are partially uncertain — not known exactly when the problem is being solved.
♠ Sources of data uncertainty:

- Prediction errors. Some of data entries (future demands, returns, etc.) do not exist when the problem is solved and hence are replaced with their forecasts.
• **Measurement errors:** Some of the data (parameters of technological devices and processes, contents associated with raw materials, etc.) cannot be measured exactly, and their true values drift around the measured “nominal” values.

• **Implementation errors:** Some of the decision variables (planned intensities of technological processes, parameters of physical devices we are designing, etc.) cannot be implemented exactly as computed. *The implementation errors are equivalent to artificial data uncertainties.*

Indeed, the impact of implementation errors $x_j \mapsto (1 + \epsilon_j)x_j + \delta_j$ on the validity of the constraint

$$a_{i1}x_1 + \ldots + a_{in}x_n \leq b_i$$

is as if there were no implementation errors, but the data of the constraint was subject to perturbations

$$a_{ij} \mapsto (1 + \epsilon_j)a_{ij}, \quad b_i \mapsto b_i - \sum_j a_{ij}\delta_j.$$
Data Uncertainty: Traditional Treatment and Dangers

♣ Traditionally,
♠ “small” (fractions of percents) data uncertainty is just ignored, the problem is solved “as it is” – with the nominal data, and the resulting nominal optimal solution is forwarded to the end user;
♠ “large” data uncertainty is assigned with a probability distribution and is treated via Stochastic Programming techniques.

♠ Fact: in many cases, even small data uncertainty can make the nominal solution heavily infeasible and thus practically meaningless.
Example: Antenna Design

[Physics:] Directional density of energy transmitted by an monochromatic antenna placed at the origin is proportional to $|D(\delta)|^2$, where the antenna’s diagram $D(\delta)$ is a complex-valued function of 3-D direction (unit 3-D vector) $\delta$.

[Physics:] For an antenna array — a complex antenna comprised of a number of antenna elements, the diagram is

$$D(\delta) = \sum_j x_j D_j(\delta) \quad (*)$$

- $D_j(\cdot)$: diagrams of elements
- $x_j$: complex weights – design parameters responsible for how the elements in the array are invoked.

Antenna Design problem: Given diagrams $D_1(\cdot), \ldots, D_k(\cdot)$ and a target diagram $D_*(\cdot)$, find the weights $x_i \in \mathbb{C}$ such that the synthesized diagram $(*)$ is as close as possible to the target diagram $D_*(\cdot)$.

When $D_j(\cdot)$, $D_*(\cdot)$, same as the weights, are real and the “closeness’ is quantified by the uniform norm on a finite grid $\Gamma$ of directions, Antenna Design becomes the LO problem

$$\min_{x \in \mathbb{R}^n, \tau} \left\{ \tau : -\tau \leq D_*(\delta) - \sum_j x_j D_j(\delta) \leq \tau \ \forall \delta \in \Gamma \right\}.$$
Example: Consider planar antenna array comprised of 10 elements (circle surrounded by 9 rings of equal areas) in the plane XY (Earth’s surface”), and our goal is to send most of the energy “up,” along the $12^\circ$ cone around the Z-axis:

- Diagram of a ring $\{z = 0, a \leq \sqrt{x^2 + y^2} \leq b\}$:
  
  $D_{a,b}(\theta) = \frac{1}{2} \int_a^b \left[ \int_0^{2\pi} r \cos(2\pi \lambda^{-1} \cos(\theta) \cos(\phi)) \, d\phi \right] \, dr$,  

- $\theta$: altitude angle  
  - $\lambda$: wavelength

10 antenna elements, equal areas, outer radius 1 m  

Diagrams of the elements vs the altitude angle $\theta$, $\lambda = 50$ cm  

- Nominal design problem:
  
  $\tau_* = \min_{x \in \mathbb{R}_{10}, \tau} \left\{ \tau : -\tau \leq D_*(\theta_i) - \sum_{j=1}^{10} x_j D_j(\theta_i) \leq \tau, \ 1 \leq i \leq 240, \ \theta_i = \frac{i \pi}{480} \right\}$

Target (blue) and nominal optimal (magenta) diagrams, 

$\tau_* = 0.0589$
**But:** The design variables are characteristics of physical devices and as such they cannot be implemented exactly as computed. What happens when there are implementation errors:

\[ x_j^{\text{fact}} = (1 + \xi_j)x_j^{\text{comp}}, \quad \xi_j \sim \text{Uniform}[-\rho, \rho] \]

with small \( \rho \)?

![Graphs showing comparison between dream and reality](image)

“Dream and reality,” nominal optimal design: samples of 100 actual diagrams (red) for different uncertainty levels. Blue: the target diagram

<table>
<thead>
<tr>
<th>Dream</th>
<th>Reality</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho = 0 )</td>
<td>( \rho = 0.0001 )</td>
</tr>
<tr>
<td>Value</td>
<td>min</td>
</tr>
<tr>
<td>( | \cdot |_\infty )-distance to target</td>
<td>0.059</td>
</tr>
<tr>
<td>Energy concentration</td>
<td>85.1%</td>
</tr>
</tbody>
</table>

Quality of nominal antenna design: dream and reality. Data over 100 samples of actuation errors per each uncertainty level \( \rho \).

♠ **Conclusion:** Nominal optimal design is completely meaningless...
Example: NETLIB Case Study.

NETLIB: a collection of LO problems for testing LO algorithms.

Constraint # 372 of the NETLIB problem PILOT4:

\[ \mathbf{a}^T \mathbf{x} \equiv -15.79081 x_{826} - 8.598819 x_{827} - 1.88789 x_{828} - 1.362417 x_{829} - 1.526049 x_{830} \\
-0.031883 x_{849} - 28.725555 x_{850} - 10.792065 x_{851} - 0.190044 x_{852} - 2.757176 x_{853} \\
-12.290832 x_{854} + 717.562256 x_{855} - 0.057865 x_{856} - 3.785417 x_{857} - 78.30661 x_{858} \\
-122.163055 x_{859} - 6.46609 x_{860} - 0.48371 x_{861} - 0.615264 x_{862} - 1.353783 x_{863} \\
-84.644257 x_{864} - 122.459045 x_{865} - 43.15593 x_{866} - 1.712592 x_{870} - 0.401597 x_{871} \\
+x_{880} - 0.946049 x_{898} - 0.946049 x_{916} \geq b \equiv 23.387405 \]

The related nonzero coordinates in the optimal solution \( \mathbf{x}^* \) of the problem as reported by CPLEX are:

\[
\begin{align*}
x_{826}^* &= 255.6112787181108 \\
x_{827}^* &= 6240.48891232100 \\
x_{828}^* &= 3624.613324098961 \\
x_{829}^* &= 18.20205065283259 \\
x_{849}^* &= 174397.0389573037 \\
x_{850}^* &= 14250.00176680900 \\
x_{851}^* &= 25910.00731692178 \\
x_{852}^* &= 104958.3199274139
\end{align*}
\]

This solution makes (C) an equality within machine precision.

Note: The coefficients in \( \mathbf{a} \), except for the coefficient 1 at \( x_{880} \), are “ugly reals” like -15.79081 or -84.644257. Ugly coefficients characterize certain technological devices and processes; as such they could hardly be known to high accuracy and coincide with the “true” data within accuracy of 3-4 digits, not more.

Question: Assuming that the ugly entries in \( \mathbf{a} \) are 0.1%-accurate approximations of the true data \( \tilde{\mathbf{a}} \), what is the effect of this uncertainty on the validity of the “true” constraint \( \tilde{\mathbf{a}}^T \mathbf{x} \geq b \) as evaluated at \( \mathbf{x}^* \)?
Answer:
• The minimum, over all 0.1\% perturbations \( a \mapsto \tilde{a} \) of ugly entries in \( a \), value of \( \tilde{a}^T x^* - b \), is \( < -104.9 \), that is, with 0.1\% perturbations of ugly coefficients, the violation of the constraint as evaluated at the nominal solution can be as large as 450\% of the right hand side!
• With independent random 0.1\%-perturbations of ugly coefficients,
  — the violation of the constraint at average is as large as 125\% of the right hand side;
  — the probability of violating the constraint by at least 150\% of the right hand side is as large as 0.18.
♣ Among 90 NETLIB problems, perturbing ugly coefficients by just 0.01\% results in violating some of the constraints, as evaluated at nominal optimal solutions,
  — by more than 50\% – in 13 problems,
  — by more than 100\% – in 6 problems.
  — by 210,000\% – in PILOT4.
Conclusion: In applications of LO, there exists a real need of a technique capable of detecting cases when data uncertainty can heavily affect the quality of the nominal solution, and in these cases to generate a “reliable” solution, one that is immunized against uncertainty. Robust Optimization is aimed at satisfying the above need.

Uncertain Linear Optimization Problems

Definition: An uncertain LO problem is a collection

$$\left\{ \min_x \{ c^T x + d : Ax \leq b \} \right\}_{(c,d,A,b) \in \mathcal{U}}$$

of LO problems (instances) \( \min_x \{ c^T x + d : Ax \leq b \} \) of common structure (i.e., with common numbers \( m \) of constraints and \( n \) of variables) with the data varying in a given uncertainty set \( \mathcal{U} \subset \mathbb{R}^{(m+1) \times (n+1)} \).

Usually we assume that the uncertainty set is parameterized, in an affine fashion, by perturbation vector \( \zeta \) varying in a given perturbation set \( \mathcal{Z} \):

$$\mathcal{U} = \left\{ \begin{bmatrix} c^T \\ A \\ d \\ b \end{bmatrix} = \begin{bmatrix} c_0^T \\ A_0 \\ d_0 \\ b_0 \end{bmatrix} + \sum_{\ell=1}^{L} \zeta_\ell \begin{bmatrix} c_\ell^T \\ A_\ell \\ d_\ell \\ b_\ell \end{bmatrix} : \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\}.$$
**Example:** When speaking about PILOT4, we tacitly used the following model of uncertainty:

Uncertainty affects only the ‘ugly” coefficients \( \{ a_{ij} : (i, j) \in J \} \) in the constraint matrix, and every one of them is allowed to run, independently of all other coefficients, through the interval

\[
[a_{ij}^n - \rho_{ij} | a_{ij}^n| , a_{ij}^n + \rho_{ij} | a_{ij}^n|]
\]

- \( a_{ij}^n \): nominal values of the data
- \( \rho_{ij} \): perturbation levels (which in the experiment were set to \( \rho = 0.001 \)).

**Perturbation set:** The box

\[
\{ \zeta = \{ \zeta_{ij} \}_{(i,j) \in J} : -\rho_{ij} \leq \zeta_{ij} \leq \rho_{ij} \}
\]

**Parameterization of the data by perturbation vector:**

\[
\begin{bmatrix}
  c^T \\
  A \\
  b
\end{bmatrix} =
\begin{bmatrix}
  [c^n]^T \\
  A^n \\
  b^n
\end{bmatrix} + \sum_{(i,j) \in J} \zeta_{ij} \begin{bmatrix}
  e_i \\
  e_j^T
\end{bmatrix}
\]
\[
\begin{align*}
\mathcal{U} &= \left\{ \min_x \left\{ c^T x + d : Ax \leq b \right\} \right\}_{(c,d,A,b) \in \mathcal{U}} \\
&= \left\{ \begin{bmatrix} c^T_0 & d_0 \\ A_0 & b_0 \end{bmatrix} + \sum_{\ell=1}^L \zeta_\ell \begin{bmatrix} c^T_\ell & d_\ell \\ A_\ell & b_\ell \end{bmatrix} : \zeta \in \mathbb{Z} \subset \mathbb{R}^L \right\}.
\end{align*}
\]

There is no universally defined notion of a “solution to a family of optimization problems,” like \((\text{LO}_\mathcal{U})\).

Consider “decision environment” as follows:

A.1. All decision variables in \((\text{LO}_\mathcal{U})\) represent “here and now” decisions; they should be assigned specific numerical values as a result of solving the problem before the actual data “reveals itself.”

A.2. The decision maker is fully responsible for consequences of the decisions to be made when, and only when, the actual data is within the prespecified uncertainty set \(\mathcal{U}\).

A.3. The constraints in \((\text{LO}_\mathcal{U})\) are hard — we cannot tolerate violations of constraints, even small ones, when the data is in \(\mathcal{U}\).
\[
\left\{ \min_x \left\{ c^T x + d : A x \leq b \right\} \right\}_{(c,d,A,b) \in \mathcal{U}} \tag{LO_\mathcal{U}}
\]

In the above decision environment, the only meaningful candidate solutions to \( (LO_\mathcal{U}) \) are the robust feasible ones.

**Definition:** \( x \in \mathbb{R}^n \) is called a robust feasible solution to \( (LO_\mathcal{U}) \), if \( x \) is feasible for all instances:

\[
A x \leq b \quad \forall (c, d, A, b) \in \mathcal{U}.
\]

Indeed, by A.1 a meaningful candidate solution should be independent of the data, i.e., it should be just a fixed vector \( x \). By A.2-3, it should satisfy the constraints, whatever be a realization of the data from \( \mathcal{U} \).

Acting in the same “worst-case-oriented” fashion, it makes sense to quantify the quality of a candidate solution \( x \) by the guaranteed (the worst, over the data from \( \mathcal{U} \)) value of the objective:

\[
\sup \{ c^T x + d : (c, d, A, b) \in \mathcal{U} \}
\]
\[
\left\{ \min_x \{ c^T x + d : Ax \leq b \} \right\}_{(c,d,A,b) \in U} \tag{LO_u}
\]

Â Now we can associate with \( (LO_U) \) the problem of finding the best, in terms of the guaranteed value of the objective, among the robust feasible solutions:

\[
\min_{t,x} \left\{ t : c^T x + d \leq t, Ax \leq b \; \forall (c, d, A, b) \in U \right\} \tag{RC}
\]

This is called the Robust Counterpart of \( (LO_u) \).

**Note:** Passing from LOs of the form

\[
\min_x \{ c^T x + d : Ax \leq b \}
\]

to their equivalents

\[
\min_{t,x} \{ t : c^T x + d \leq t, Ax \leq b \}
\]

we always may assume that the objective is certain, and the RC respects this equivalence.

⇒ We lose nothing by assuming the objective in \( (LO_u) \) certain, in which case we can think of \( U \) as of the set in the space \( \mathbb{R}^{m \times (n+1)} \) of the \( [A, b] \)-data, and the RC reads

\[
\min_x \left\{ c^T x : Ax \leq b \; \forall [A, b] \in U \right\}. \tag{RC}
\]
\[
\left\{ \min_x \{ c^T x : Ax \leq b \} \right\}_{(A,b) \in U} \quad (LO_{\mathcal{U}})
\]
\[
\downarrow
\]
\[
\min_x \{ c^T x : Ax \leq b \ \forall [A, b] \in \mathcal{U} \} \quad (RC)
\]

\large{♣ Fact I:} The RC of uncertain LO with certain objective is a purely constraint-wise construction: when building the RC, we replace every constraint \( a_i^T x \leq b_i \) of the instances with its RC
\[
a_i^T x \leq b_i \ \forall [a_i^T, b_i] \in \mathcal{U}_i
\]
where \( \mathcal{U}_i \) is the projection of the uncertainty set \( \mathcal{U} \) on the space of data \([a_i^T, b_i]\) of \(i\)-th constraint.

\large{♣ Fact II:} The RC remains intact when extending the uncertainty set \( \mathcal{U} \) to its closed convex hull.
When \((LO_{\mathcal{U}})\) has certain objective, the RC remains intact when extending \(\mathcal{U}\) to the direct product of closed convex hulls of \(\mathcal{U}_i\). Thus, the transformation
\[
\mathcal{U} \mapsto \mathcal{U}^+ = [\text{cl Conv}(\mathcal{U}_1)] \times \ldots \times [\text{cl Conv}(\mathcal{U}_m)]
\]
keeps the RC intact.

\large{♦ From now on, we always assume uncertainty set \( \mathcal{U} \) convex, and perturbation set \( \mathcal{Z} \) – convex and closed.
\[
\min \left\{ c^T x : Ax \leq b \right\}_{[A,b] \in U} \quad (\text{LO}_U)
\]
\[
\downarrow
\]
\[
\min_x \left\{ c^T x : Ax \leq b \quad \forall [A,b] \in U \right\} \quad (\text{RC})
\]

♣ The central questions associated with the concept of RC are:

**A.** What is the “computational status” of the RC? When is it possible to process the RC efficiently?
— to be addressed in-depth below.

**B.** How to come-up with meaningful uncertainty sets?
— modeling issue to be partly addressed in the sequel.
\[
\min_x \{ c^T x : Ax \leq b \ \forall [A, b] \in U \} \quad \text{(RC)}
\]

\begin{itemize}
\item \textbf{Potentially bad news:} The RC is a semi-infinite optimization problem (finitely many variables, infinitely many constraints) and as such can be computationally tractable.
\end{itemize}

\begin{itemize}
\item \textbf{Example:} Consider an “essentially linear” semi-infinite constraint
\[
\| Px - p \|_1 \leq 1, \ \forall [P, p] \in U
\]
\[
U = \{ [P_*, p] : p = B\zeta, \|\zeta\|_2 \leq 1 \}
\]
To check whether \( x = 0 \) is robust feasible is the same as to check whether
\[
\max_{\zeta : \|\zeta\|_2 \leq 1} \| B\zeta \|_1 \leq 1. \quad (!)
\]

(!) is equivalent to
\[
1 \geq \max_{\|\zeta\|_2 \leq 1} \| B\zeta \|_1 = \max_{z : \|z\|_\infty \leq 1, \zeta : \|\zeta\|_2 \leq 1} z^T B\zeta
\]
\[
= \max_{z : \|z\|_\infty \leq 1} \max_{\zeta : \|\zeta\|_2 \leq 1} 1 \cdot \zeta^T [B^T z] = \sqrt{\max_{z : \|z\|_\infty \leq 1} z^T [BB^T] z}
\]

Since \( BB^T \) can be an arbitrary symmetric positive semidefinite matrix, and finding the maximum of a nonnegative quadratic form over the box \( \{ z : \|z\|_\infty \leq 1 \} \) is NP-hard, even when relative accuracy like 4\% is sought, \textit{checking (!) is heavily computationally intractable.}
\[ \min_x \{ c^T x : Ax \leq b \ \forall [A, b] \in \mathcal{U} \} \quad (\text{RC}) \]

♣ **Good news:** The RC of an uncertain LO problem is computationally tractable, provided the uncertainty set \( \mathcal{U} \) is so.

Explanation, I: The RC can be written down as the optimization problem

\[
\min_x \{ c^T x : f_i(x) \leq 0, \ i = 1, \ldots, m \} \\
 f_i(x) = \sup_{[A,b] \in \mathcal{U}} [a_i^T x - b_i]
\]

- The functions \( f_i(x) \) are convex (due to their origin) and efficiently computable (as maxima of affine functions over computationally tractable convex sets).
- Thus, the RC is a Convex Programming program with efficiently computable objective and constraints, and problems of this type are efficiently solvable.
The above “reasoning” refers to the notions of computationally tractable problem/convex set and on the fact that maximizing linear objective over a computationally tractable convex set, in particular, a convex set given by finitely many efficiently computable convex constraints, is a computationally tractable problem. While these notions and results can be rigorously defined and justified, it makes sense to present a somehow restricted “practical” version of them, highly instructive by its own rights and not requiring tedious and lengthy excursions to the complexity theory of continuous optimization.
Recalling that the RC is a “constraint-wise” construction, all we need is to reformulate in a tractable form a single semi-infinite constraint

\[ \forall \alpha = [a; b] \in \{ \alpha_0 + A\zeta : \zeta \in \mathcal{Z} \} \subset \mathbb{R}^{n+1} : \alpha^T[x; 1] \equiv a^T x + b \leq 0. \]  

Consider several instructive cases when tractable reformulation of (\*\*) is easy – does not require any theory.

1. **Scenario uncertainty** \( \mathcal{Z} = \text{Conv}\{\zeta^1, ..., \zeta^N\} \). Setting \( \alpha^j = \alpha_0 + A\zeta^j \), \( 1 \leq j \leq N \), we get

\[ \mathcal{U} = \text{Conv}\{\alpha^1, ..., \alpha^N\} \]

and therefore

\[ (\*) \Leftrightarrow \{ \alpha^j[x; 1] \leq 0, 1 \leq j \leq N \} \]

2. **\( \| \cdot \|_p \)-uncertainty** \( \mathcal{Z} = \{ \zeta \in \mathbb{R}^L : \| \zeta \|_p \leq 1 \} \). We have

\[
\begin{align*}
\alpha^T[x; 1] & \leq 1 \forall \alpha \in \mathcal{U} \\
\Leftrightarrow \ [\alpha_0 + A\zeta]^T[x; 1] & \leq 0 \forall (\zeta : \| \zeta \|_p \leq 1) \\
\Leftrightarrow \ [\alpha_0]^T[x; 1] + \max_{\| \zeta \|_p \leq 1} \zeta^T[A^T[x; 1]] & \leq 0 \\
\Leftrightarrow \ [\alpha_0]^T[x; 1] + \|A^T[x; 1]\|_{p^*} & \leq 0, \frac{1}{p} + \frac{1}{p^*} = 1
\end{align*}
\]
\[ \forall \alpha = [a; b] \in \{ \alpha = \alpha_0 + A\zeta : \zeta \in \mathcal{Z} \} \subset \mathbb{R}^{n+1} : \]
\[ \alpha^T [x; 1] := a^T x + b \leq 0. \quad (*) \]

3. Intersection of simple perturbation sets: \( \mathcal{Z} = \bigcap_{i=1}^k \mathcal{Z}_i \). Let \( \mathcal{Z}_i, 0 \in \mathcal{Z}_i, 1 \leq i \leq k \) be convex compact sets such that \( \bigcap_{i=1}^k \text{int}\mathcal{Z}_i \neq \emptyset \).

Fact from Convex Analysis: For \( \mathcal{Z}, \mathcal{Z}_i \) as above,
\[ \max_{\zeta \in \mathcal{Z}} \beta^T \zeta = \min_{\beta_1, \ldots, \beta_k} \sum_{i=1}^k \max_{\zeta \in \mathcal{Z}_i} \beta_i^T \zeta. \]

Therefore,
\[ \alpha^T [x; 1] \leq 0 \quad \forall \alpha \in \mathcal{U} \]
\[ \iff \alpha_0^T [x; 1] + [A\zeta]^T [x; 1] \leq 0 \quad \forall \zeta \in \mathcal{Z} \]
\[ \iff \alpha_0^T [x; 1] + \max_{\zeta \in \mathcal{Z}} \zeta^T [A^T [x; 1]] \leq 0 \]
\[ \iff \exists \beta_1, \ldots, \beta_k : \]
\[ \left\{ \begin{array}{l}
\beta_1 + \ldots + \beta_k = A^T [x; 1] \\
\alpha_0^T [x; 1] + \sum_{i=1}^k \max_{\zeta \in \mathcal{Z}_i} \beta_i^T \zeta \leq 0
\end{array} \right. \quad (a) \quad (b) \]

Thus, (*) is represented by the system
\[ \left\{ \begin{array}{l}
\beta_1 + \ldots + \beta_k = A^T [x; 1] \\
\alpha_0^T [x; 1] + \sum_{i=1}^k \max_{\zeta \in \mathcal{Z}_i} \beta_i^T \zeta \leq 0
\end{array} \right. \quad (S) \]

of constraints in variables \( x, \beta_1, \ldots, \beta_k \), meaning that \( x \) can be extended to a feasible solution of (S) if and only if \( x \) is feasible for (*).
When \( \mathcal{Z} = \bigcap_{i=1}^{k} \mathcal{Z}_i, \ 0 \in \mathcal{Z}, \ \bigcap_i \text{int} \mathcal{Z}_i \neq \emptyset \), the system
\[
\begin{align*}
\begin{cases}
\beta_1 + \ldots + \beta_k &= \mathbf{A}^T[x; 1] \quad (a) \\
\alpha_0^T[x; 1] + \sum_{i=1}^{k} \max_{\zeta \in \mathcal{Z}_i} \beta_i^T \zeta &\leq 0 \quad (b)
\end{cases}
\end{align*}
\]
of convex constraints in variables \( x, \beta_1, \ldots, \beta_k \) represents the semi-infinite constraint
\[
\alpha^T[x; 1] \leq 0 \ \forall \alpha \in \{ \alpha_0 + \mathbf{A} \zeta : \zeta \in \mathcal{Z} \}
\]

**Note:** When \( \mathcal{Z}_i \) are simple, so that the convex functions \( f_i(\beta_i) = \max_{\zeta \in \mathcal{Z}_i} \beta_i^T \zeta \) are available in closed analytic form, \( (S) \) is a system of explicitly given convex constraints.

**Example:** Ball-Box-Budgeted uncertainty
\[
\mathcal{Z} = \{ \zeta : \|\zeta\|_\infty \leq \Omega_\infty \} \cap \{ \zeta : \|\zeta\|_2 \leq \Omega_2 \} \cap \{ \zeta : \|\zeta\|_1 \leq \Omega_1 \}.
\]
Here
\[
\begin{align*}
f_1(\beta) &= \max_{\zeta : \|\zeta\|_\infty \leq \Omega_\infty} \beta^T \zeta = \Omega_\infty \|\beta\|_1, \\
f_2(\beta) &= \max_{\zeta : \|\zeta\|_2 \leq \Omega_2} \beta^T \zeta = \Omega_2 \|\beta\|_2, \\
f_3(\beta) &= \max_{\zeta : \|\zeta\|_1 \leq \Omega_1} \beta^T \zeta = \Omega_1 \|\beta\|_\infty,
\end{align*}
\]
and thus \( (S) \) is equivalent to the system of convex constraints
\[
\begin{align*}
\begin{cases}
\beta_1 + \beta_2 + \beta_3 &= \mathbf{A}^T[x; 1] \\
\alpha_0^T[x; 1] + \Omega_\infty \|\beta_1\|_1 + \Omega_2 \|\beta_2\|_2 + \Omega_1 \|\beta_3\|_\infty &\leq 0
\end{cases}
\end{align*}
\]
General Well-Structured Case

Definition. Let us say that a set $\mathcal{X} \subset \mathbb{R}^N$ is well-structured, if it admits a well-structured representation — a representation of the form

$$\mathcal{X} = \left\{ x \in \mathbb{R}^N : \exists u \in \mathbb{R}^M : \begin{cases} A_0 x + B_0 u + c_0 = 0 \\ A_1 x + B_1 u + c_1 \in K_1 \\ \ldots \\ A_K x + B_K u + c_k \in K_K \end{cases} \right\},$$

where $K_k$, for every $k \leq K$, is a simple cone, specifically,
— either nonnegative orthant $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x \geq 0\}$, $m = m_k$,
— or a Lorentz cone $L^m = \{x \in \mathbb{R}^m : x_m \geq \sqrt{x_1^2 + \ldots + x_{m-1}^2}\}$, $m = m_k$,
— or a Semidefinite cone $S^m_+$ — the cone of positive semidefinite matrices in the space $S^m$ of real symmetric $m \times m$ matrices, $m = m_k$. 

22
**Example 1:** The set \( \mathcal{X} = \{ x \in \mathbb{R}^N : \| x \|_1 \leq 1 \} \) admits polyhedral representation

\[
\mathcal{X} = \{ x \in \mathbb{R}^N : \exists u \in \mathbb{R}^N : -u_i \leq x_i \leq u_i, \sum_i u_i \leq 1 \}
\]

\[
= \left\{ x \in \mathbb{R}^n : \exists u \in \mathbb{R}^N : A_1 x + B_1 u + c_1 \equiv \begin{bmatrix}
  u_1 - x_1 \\
  u_1 + x_1 \\
  \vdots \\
  u_N - x_N \\
  u_N + x_N \\
  1 - \sum_i u_i
\end{bmatrix} \in \mathbb{R}_+^{2N+1} \right\}
\]

**Example 2:** The set \( \mathcal{X} = \{ x \in \mathbb{R}_+^4 : x_1 x_2 x_3 x_4 \geq 1 \} \) admits conic quadratic representation

\[
\mathcal{X} = \left\{ x \in \mathbb{R}^4 : \exists u \in \mathbb{R}^3 : \begin{bmatrix} 0 \leq u_1 \leq \sqrt{x_1 x_2} \\ 0 \leq u_2 \leq \sqrt{x_3 x_4} \\ 1 \leq u_3 \leq \sqrt{u_1 u_2} \end{bmatrix} \right\}
\]

\[
= \left\{ x \in \mathbb{R}^n : \exists u \in \mathbb{R}^3 : \begin{bmatrix} x_1; x_2; x_3; x_4; u_1; u_2; u_3 - 1 \end{bmatrix} \in \mathbb{R}_+^7 \\
[2u_1; x_1 - x_2; x_1 + x_2] \in \mathbb{L}_3^3 \\
[2u_2; x_3 - x_4; x_3 + x_4] \in \mathbb{L}_3^3 \\
[2u_3; u_1 - u_2; u_1 + u_2] \in \mathbb{L}_3^3 \right\}
\]

**Example 3:** The set \( \mathcal{X} \) of \( m \times n \) matrices \( X \) with nuclear norm (sum of singular values) \( \leq 1 \) admits semidefinite representation

\[
\mathcal{X} = \left\{ X \in \mathbb{R}^{m \times n} : \exists u = (U \in \mathbb{S}^m, V \in \mathbb{S}^n) : \right\}
\]

\[
\text{Tr}(U) + \text{Tr}(V) \leq 2 \\
\begin{bmatrix} U & X \\ X^T & V \end{bmatrix} \succeq 0 \right\}
\]
\[ \mathcal{X} = \left\{ x \in \mathbb{R}^N : \exists u \in \mathbb{R}^M : \begin{cases} A_0x + B_0u + c_0 = 0 \\ A_1x + B_1u + c_1 \in K_1 \\ \vdots \\ A_Kx + B_ku + c_k \in K_K \end{cases} \right\}, \quad (\ast) \]

\[ \min_{x,u} \begin{pmatrix} c^T x : \begin{cases} A_0x + B_0u + c_0 = 0 \\ A_1x + B_1u + c_1 \in K_1 \\ \vdots \\ A_Kx + B_ku + c_k \in K_K \end{cases} \end{pmatrix}, \]

\[ \begin{array}{c}
\textbf{Good news on well-structured representations:}
\end{array} \]

- **Computational tractability:** Minimizing a linear objective over a set given by (\ast) reduces to solving a well-structured conic program

\[ \begin{array}{c}
\text{and thus can be done in a theoretically (and to some extent — also practically) efficient manner by polynomial time interior point algorithms.}
\end{array} \]

- **Extremely powerful expressive abilities:** w.-s.r.’s admit a simple fully algorithmic calculus which makes it easy to build a w.-s.r. for the result of a convexity-preserving operation with convex sets (like taking intersections, direct sums, affine images, inverse affine images, polars, etc.) via w.-s.r.’s of the operands.

As a result, for all practical purposes, all computationally tractable convex sets arising in Optimization admit explicit w.-s.r.’s.
The RC Tractability Theorem: Let the perturbation set $\mathcal{Z}$ of a semi-infinite linear inequality

$$\alpha^T [x; 1] \leq 0 \quad \forall \alpha \in \{ \alpha_0 + A \zeta : \zeta \in \mathcal{Z} \}$$  \hspace{1cm} (\ast)$$

be nonempty and be given by w.-s.r.

$$\mathcal{Z} = \left\{ \zeta \in \mathbb{R}^L : \exists u \in \mathbb{R}^M : \begin{cases} A_0 x + B_0 u + c_0 = 0 \\ A_1 x + B_1 u + c_1 \in \mathbf{K}_1 \\ \ldots \\ A_K x + B_k u + c_k \in \mathbf{K}_K \end{cases} \right\} \hspace{1cm} (\dagger)$$

When not all the cones $\mathbf{K}_k$ are nonnegative orthants, assume that (\dagger) is strictly feasible, that is, there exist $\bar{x}$ and $\bar{u}$ such that

$$A_0 \bar{x} + B_0 \bar{u} + c_0 = 0 \land A_k \bar{x} + B_k \bar{u} + c_k \in \text{int} \mathbf{K}_k, \; 1 \leq k \leq K.$$  

Then the feasible set $\mathcal{X}$ of (\ast) admits an explicit w.-s.r., specifically,

$$\mathcal{X} = \left\{ x : \exists z = [z^0 ; \ldots ; z^K] : \begin{cases} \sum_{k=0}^K A_k^* z^k + A^T [x; 1] = 0 \\ \sum_{k=0}^K B_k^* z^k = 0 \\ \alpha_0^T [x; 1] + \sum_{k=0}^K \langle z^k, c_k \rangle \leq 0 \\ z^k \in \mathbf{K}_k, \; 1 \leq k \leq K \end{cases} \right\}$$

Here for a linear map $e \mapsto Be$ from a Euclidean space $(E, \langle \cdot , \cdot \rangle_E)$ to a Euclidean space $(F, \langle \cdot , \cdot \rangle_F)$ the adjoint map $f \mapsto B^* f : F \to E$ is given by

$$\langle f , Be \rangle_F \equiv \langle B^* f , e \rangle_E$$
Proof heavily utilizes the Conic Duality Theorem which answers the following question:

♣ Consider a conic program

\[
\text{Opt}(P) = \min_y \begin{cases} 
\langle c, y \rangle : \\
A_0y - b_0 = 0 \\
A_ky - b_k \in K_k, \\
1 \leq k \leq K 
\end{cases},
\]  

(P)

where \( K_k \) are cones (closed, convex, pointed and with a nonempty interior) in Euclidean spaces \( E_k \), \( 1 \leq k \leq K \).

How to bound from below, in a systematic way, the optimal value of the program?

♠ Consider an approach as follows. Let \( K^*_k = \{ u \in E_k : \langle u, v \rangle \geq 0 \, \forall v \in K_k \} \) be the cones dual to \( K_k \). Let us choose \( z^0 \in \mathbb{R}^{\text{dim} b_0} \) and \( z^k \in K^*_k, 1 \leq k \leq K \), and let \( y \) be feasible for (P). By feasibility, we have

\[
\langle z^k, A_ky - b_k \rangle \geq 0, \quad 0 \leq k \leq K,
\]

or, which is the same,

\[
\langle A^*_k z^k, y \rangle \geq \langle z^k, b_k \rangle, \quad 0 \leq k \leq K.
\]

Summing up, we get

\[
\langle \sum_{k=0}^K A^*_k z^k, y \rangle \geq \sum_{k=0}^K \langle z^k, b_k \rangle.
\]
\[
\text{Opt}(P) = \min_y \left\{ \langle c, y \rangle : \begin{array}{l}
A_0 y - b_0 = 0 \\
A_k y - b_k \in K_k, \\
1 \leq k \leq K
\end{array} \right\}, \quad (P)
\]

**Intermediate summary:** Whenever \( z^0 \in \mathbb{R}^\text{dim} b_0 \) and \( z^k \in K_k^*, \ 1 \leq k \leq K, \) every feasible solution \( y \) of \( (P) \) satisfies the inequality
\[
\langle \sum_{k=0}^K A_k^* z^k, y \rangle \geq \sum_{k=0}^K \langle z^k, b_k \rangle. \quad (*)
\]

**Conclusion:** When the left hand side in \( (*) \) is identically in \( y \in \mathbb{R}^N \) equal to \( \langle c, y \rangle \), the right hand side in \( (*) \) is a lower bound on \( \text{Opt}(P) \). In other words, The optimal value \( \text{Opt}(D) \) in the conic dual of \( (P) \), that is, in the problem
\[
\text{Opt}(D) = \max_{\{z^k\}} \left\{ \sum_{k=0}^K \langle z^k, b_k \rangle : \begin{array}{l}
z^k \in K_k^*, \\
1 \leq k \leq K \\
\sum_{k=0}^K A_k^* z^k = c
\end{array} \right\} \quad (D)
\]

is a lower bound on \( \text{Opt}(P) \). [“Weak Duality”]

♣ **Conic Duality Theorem:** If \( (P) \) is strictly feasible and below bounded, then \( (D) \) is solvable, and \( \text{Opt}(P) = \text{Opt}(D) \).

**Note:** When \( K_k = \mathbb{R}^{mk}_+ \) for all \( k \), “strict feasibility” can be weakened to “feasibility.”
\[ \alpha^T[x; 1] \leq 0 \quad \forall \alpha \in \{ \alpha_0 + A\zeta : \zeta \in \mathbb{Z} \} \]

\[ Z = \left\{ \zeta \in \mathbb{R}^L : \exists u \in \mathbb{R}^M : \begin{cases} \ A_0x + B_0u + c_0 = 0 \\ A_1x + B_1u + c_1 \in K_1 \\ \vdots \\ A_Kx + B_Ku + c_k \in K_K \end{cases} \right\} \]

Observe that \( x \) is feasible for (*) iff

\[ \text{Opt}(P) := \min_{\zeta \in Z} \{-A\zeta^T[x; 1]\} \geq \alpha_0^T[x; 1], \]

or, which is the same, iff

\[ \text{Opt}(P) := \min_{\zeta, u} \{-A^T[x; 1]\} \zeta : \begin{cases} \ A_0x + B_0u + c_0 = 0 \\ A_1x + B_1u + c_1 \in K_1 \\ \vdots \\ A_Kx + B_Ku + c_k \in K_K \end{cases} \geq \alpha_0^T[x; 1] \]

By CDT, and noting that \( K^*_k = K_k \) for our cones, this is the case iff the problem

\[ \max_{\{z^0, \ldots, z^K\}} \left\{ -\sum_{k=0}^{L} \langle z^k, c_k \rangle : \begin{cases} \sum_{k=0}^{K} A_k^*z^k = -A^T[x; 1] \\ \sum_{k=0}^{K} B_k^*z^k = 0 \\ z^k \in K_k, 1 \leq k \leq K \end{cases} \right\} \]

has a solution with the value of the objective \( \geq \alpha_0^T[x; 1] \).
• Thus, $x$ is feasible for $(\ast)$ iff there exists $z = [z^0; \ldots; z^K]$ such that

\[
\begin{align*}
\alpha_0^T [x; 1] + \sum_{k=0}^K \langle z^k, c_k \rangle & \leq 0 \\
\sum_{k=0}^K A_k^* z^k + \mathcal{A}^T [x; 1] & = 0 \\
\sum_{k=0}^K B_k^* z^k & = 0 \\
z^k & \in \mathbf{K}_k, \ 1 \leq k \leq K
\end{align*}
\]
How it Works: Antenna Design

\[
\begin{align*}
\min_{\tau, x} \left\{ \tau : \tau - D_*(\theta_i) - \sum_{j=1}^{10} x_j D_j(\theta_i) &\leq \tau, 1 \leq i \leq I \right\} \\
x_j &\mapsto (1 + \zeta_j) x_j, \quad -\rho \leq \zeta_j \leq \rho \\
\Rightarrow \min_{\tau, x} \left\{ \tau : \begin{array}{l}
D_*(\theta_i) - \sum_j x_j D_j(\theta_i) - \rho \sum_j |x_j| |D_j(\theta_i)| \geq -\tau \\
D_*(\theta_i) - \sum_j x_j D_j(\theta_i) + \rho \sum_j |x_j| |D_j(\theta_i)| \leq \tau
\end{array}, 1 \leq i \leq I \right\} (RC)
\end{align*}
\]

\[\mathbf{\dag} \text{ Solving (RC) at uncertainty level } \rho = 0.01, \text{ we arrive at robust design. The robust optimal value is } 0.0815 \text{ (39\% more than the nominal optimal value 0.0589).}\]

“Dream and reality,” robust optimal design: samples of 100 actual diagrams (red) for different uncertainty levels. Blue: the target diagram.

<table>
<thead>
<tr>
<th>Reality</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho = 0.01$</td>
<td>$\rho = 0.05$</td>
<td>$\rho = 0.1$</td>
<td></td>
</tr>
<tr>
<td>$| \cdot |_\infty$-distance to target</td>
<td>min</td>
<td>mean</td>
<td>max</td>
</tr>
<tr>
<td>0.075</td>
<td>0.078</td>
<td>0.081</td>
<td>0.077</td>
</tr>
<tr>
<td>Energy concentration</td>
<td>70.3%</td>
<td>72.3%</td>
<td>73.8%</td>
</tr>
</tbody>
</table>

Robust optimal design, data over 100 samples of actuation errors per each uncertainty level $\rho$. For nominal design with $\rho = 0.001$, the average $\| \cdot \|_\infty$-distance to target is 56.8, and energy concentration is 16.5%.
How it Works: NETLIB Case Study

At uncertainty level $\rho = 0.001$, the RCs of all 90 NETLIB problems are feasible, and the robust optimal values of all problems are within 1% of their nominal optimal values.