Optimization under uncertainty

Motivation: Recall the standard form LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Data of the problem: (c, A, b)

$c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$

* so far, assumed deterministic
* in reality, data are often uncertain

Example: Manufacturing problem

c_j = cost of j^{th} activity

a_{ij} = amount of i^{th} commodity per unit of j^{th} activity

b_i = output requirement of i^{th} commodity

$\rightarrow$ c_j, a_{ij} could be uncertain

How do we incorporate data uncertainty in the model?

Consider the uncertainty as a perturbation by a vector $\xi \in \mathbb{R}^l$

One idea:

$$\begin{aligned} \min \quad & c(\xi)^T x \\ \text{s.t.} \quad & A(\xi) x = b(\xi) \qquad \text{———} (\Delta) \\ & x \geq 0 \end{aligned}$$

As it stands, this is not a well-posed formulation.
What should be the interpretation of this?

Stochastic Optimization point-of-view:

ξ is a random vector with distribution $\mathbb{P}$ (not necessarily known)

Then, one way of interpreting (Δ) is

$$\begin{aligned} \min \quad & \mathbb{E}_{\mathbb{P}}[c(\xi)^T x] \\ \text{s.t.} \quad & \Pr_{\mathbb{P}}[A(\xi) x = b(\xi)] \geq 1 - \delta \qquad \text{———} (SLP) \\ & x \geq 0 \end{aligned}$$

Here, $\delta \in (0,1)$ is a parameter. The constraint is known as a probabilistic or chance constraint.

* Difficulties:
  - Hard to evaluate $\mathbb{E}_{\mathbb{P}}$ or $\Pr_{\mathbb{P}}$ in general, even if $\mathbb{P}$ is known

    $\mathbb{E}_{\mathbb{P}}[c(\xi)^T x] = \int_{\mathbb{R}^l} c(\xi)^T x \, d\mathbb{P}$ multidimensional integral

  - distribution $\mathbb{P}$ itself is not known

Robust Optimization point-of-view:

ξ belongs to a (bounded) set $\mathcal{U} \subseteq \mathbb{R}^l$, not necessarily random

$$\begin{array}{lll} \min & t & \\ \text{s.t} & \left.\begin{array}{l} c(\xi)^T x \le t \\ A(\xi) x = b(\xi) \end{array}\right\} & \forall \xi \in \mathcal{U} \qquad \text{— (RLP)} \\ & x \ge 0. & \end{array}$$

Thus, the Solution (x^*, t^*) to (RLP) is <u>robust</u> against all possible $\xi \in \mathcal{U}$.

* Difficulties:

- How to choose $\mathcal{U}$? Can the resulting problem be solved?
- Solution may be too conservative

<u>Distributionally Robust Optimization point-of-view:</u>

ξ is a random vector with distribution $\mathbb{P}^*$, and though $\mathbb{P}^*$ is not known exactly, we have some information about $\mathbb{P}^*$ (e.g., through samples of $\mathbb{P}^*$)

* Consider an uncertainty set of distributions $\mathcal{P}$:

e.g.: $\inf_{\mathbb{P} \in \mathcal{P}} \Pr_{\mathbb{P}} (A(\xi) x = b(\xi)) \ge 1 - \delta$

* Difficulties

- How to choose $\mathcal{P}$? Can the resulting problem be solved?

— Focus of this course: (Distributionally) robust optimization

* In particular, how to choose the uncertainty set $\mathcal{U}$ (or $\mathcal{P}$) and how does it affect solvability?

<u>Example</u>: Robust linear constraint

For simplicity, consider a single linear constraint subject to uncertainty:

$$\text{(RLC)} \qquad a^T x \le b \qquad \forall (a,b) = (a^0, b_0) + \sum_{j=1}^{l} \xi_j (a^j, b_j), \quad \xi \in \mathcal{U} \subseteq \mathbb{R}^l$$

where $a^0, \ldots, a^l \in \mathbb{R}^n$ and $b_0, \ldots, b_l \in \mathbb{R}$ are given. (Interpretation?)

* A common choice of $\mathcal{U}$: Some norm ball

<u>1°</u>: $\mathcal{U} = \{ y \in \mathbb{R}^l : \|y\|_\infty \le 1 \}$. Then,

$$\text{(RLC)} \iff (a^0)^T x + \sum_{j=1}^{l} \xi_j (a^j)^T x \le b_0 + \sum_{j=1}^{l} \xi_j b_j \qquad \forall \|\xi\|_\infty \le 1$$

$$\iff \sum_{j=1}^{l} \xi_j \left((a^j)^T x - b_j \right) \le b_0 - (a^0)^T x \qquad \forall \|\xi\|_\infty \le 1$$

not a finite intersection of linear constraints

$$\Leftrightarrow \quad \sum_{j=1}^{\ell} \xi_j \left((a^j)^T x - b_j \right) \le b_0 - (a^0)^T x \quad \forall \|\xi\|_\infty \le 1$$

intersection of linear constraints

$$\Leftrightarrow \quad \max_{\|\xi\|_\infty \le 1} \sum_{j=1}^{\ell} \xi_j \left((a^j)^T x - b_j \right) \le b_0 - (a^0)^T x$$

$$\Leftrightarrow \quad \sum_{j=1}^{\ell} \left| (a^j)^T x - b_j \right| \le b_0 - (a^0)^T x \longleftarrow$$

can be converted into a finite system of linear constraints $\Rightarrow$ still an LP

<u>2°</u>: $\mathcal{U} = \{ y \in \mathbb{R}^\ell : \|y\|_2 \le 1 \}$. Then,

$$(\text{RLC}) \Leftrightarrow \max_{\|\xi\|_2 \le 1} \sum_{j=1}^{\ell} \xi_j \left((a^j)^T x - b_j \right) \le b_0 - (a^0)^T x$$

$$\Leftrightarrow \left[\sum_{j=1}^{\ell} \left((a^j)^T x - b_j \right)^2 \right]^{1/2} \le b_0 - (a^0)^T x.$$

$\leftarrow$ This is no longer a linear constraint. Still, it is convex.

* Can this be efficiently solved?
* So far, we seem to be lucky in that the $\max_{\xi \in \mathcal{U}} [\ldots]$ has a closed form. In general, we do not expect this to be the case.

<u>3°</u> $\mathcal{U} = \{ y \in \mathbb{R}^\ell : Py \le q \}$ (a polyhedron)

$$(\text{RLC}) \Leftrightarrow \max_{\xi \in \mathcal{U}} \sum_{j=1}^{\ell} \xi_j \left((a^j)^T x - b_j \right) \le b_0 - (a^0)^T x$$

$$\Leftrightarrow \; ??$$

Perhaps LP duality can be applied here.

<u>Lessons learned</u>

* Complexity of the robust constraint depends on $\mathcal{U}$
* duality theory helpful in reformulating robust constraints into finite system of constraints
* even for robust linear constraints, their reformulation may not be linear $\longrightarrow$ need nonlinear optimization techniques

Recall again the standard form LP:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & \boxed{x \geq 0} \end{aligned}$$

This is comparing two vectors and requires definition.

Q: How to extend this model to incorporate nonlinearities?

Algebraic View

" $\geq$ " defines a "good" order

partial order (偏序關係)

1) (Reflexivity 自反性) $u \geq u$
2) (Anti-Symmetry 反對稱性) $\left.\begin{matrix} u \geq v \\ v \geq u \end{matrix}\right\} \Rightarrow u = v$
3) (Transitivity 遞移性) $\left.\begin{matrix} u \geq v \\ v \geq w \end{matrix}\right\} \Rightarrow u \geq w$

Also,

4) (Homogeneity 齊次性) $\left.\begin{matrix} u \geq v \\ \alpha \geq 0 \end{matrix}\right\} \Rightarrow \alpha u \geq \alpha v$
5) (Additivity 可加性) $\left.\begin{matrix} u \geq v \\ w \geq z \end{matrix}\right\} \Rightarrow u + w \geq v + z$

These allow us to prove

* Farkas lemma: certifying (in)feasibility of linear systems
* Strong duality: certifying optimality of a pair of primal-dual LPs

Geometric View

Consider $K = \{x \in \mathbb{R}^n : x_i \geq 0 \ \forall i\} = \mathbb{R}^n_+$.

This is closed and has $0 \in K$ and $\text{int}(K) = \mathbb{R}^n_{++} \neq \phi$.

Moreover, it is a pointed cone:

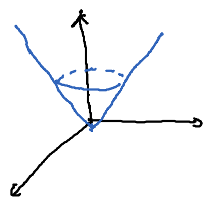
1) $K \neq \phi$; $u, v \in K \Rightarrow u + v \in K$
2) $u \in K$, $\alpha > 0 \Rightarrow \alpha u \in K$
3) $u, -u \in K \Rightarrow u = 0$.

This gives an intuitive understanding of K.

Q: Is " $\geq$ " the only relation satisfying (1)–(5)? Is $\mathbb{R}^n_+$ the only closed pointed cone containing the origin and having non-empty interior?

A: Interestingly, no!

1°. $K = Q^{n+1} = \{(t, x) \in \mathbb{R}^n \times \mathbb{R} : t \geq \|x\|_2\}$ Second-order/Lorentz cone



Exercise: Verify Q^{n+1} is closed pointed cone with $0 \in Q^{n+1}$ and $\text{int}(Q^{n+1}) \neq \phi$. Determine $\text{int}(Q^{n+1})$.

Exercise: Verify that Q^{n+1} is not polyhedral.

Consider the order " $\succeq_{Q^{n+1}}$ " defined by

$$(s, x) \succeq_{Q^{n+1}} (t, y) \iff (s - t, x - y) \in Q^{n+1}$$

Exercise: Verify that " $\succeq_{Q^{n+1}}$ " is a good order.

2°: $K = S^n_+ = \{Y \in S^n : u^T Y u \geq 0 \ \forall u\}$ Semidefinite Cone

Exercise: Verify S^n_+ is a closed pointed cone with $0 \in S^n_+$ and $\text{int}(S^n_+) \neq \phi$. Determine $\text{int}(S^n_+)$.

Consider the order "$\succeq_{S^n_+}$" defined by

$$X \succeq_{S^n_+} Y \iff X - Y \in S^n_+$$

Exercise: Verify that "$\succeq_{S^n_+}$" is a good order.

With a closed pointed cone K containing 0 and having non-empty interior, we can formulate the following problem, known as conic LP:

$$(P) \quad \begin{array}{ll} \inf & \langle c, x \rangle \\ \text{s.t.} & \langle a_i, x \rangle = b_i, \\ & x \in K \end{array}$$

Here, $\langle \cdot, \cdot \rangle$ is an inner product on some Euclidean space.

Similar to the LP dual, the dual of the conic LP is

$$(D) \quad \begin{array}{ll} \sup & b^T y \\ \text{s.t} & c - \sum_i y_i a_i \in K^*, \end{array}$$

Observe $y \mapsto c - \sum_i y_i a_i$ is affine in y. Hence, (D) can be described as optimizing a linear function s.t. conic constraint on affine map of y

where $K^* = \{ y : \langle x, y \rangle \geq 0 \ \forall x \in K \}$ is the dual cone of K.

Proposition: If K is a closed pointed cone with non-empty interior, so is K^*.

Thus, (P) and (D) have the same nature.

Exercise: Show that $\mathbb{R}^n_+$, Q^{n+1}, S^n_+ are self-dual; i.e, $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$ and so on.

Examples:

(1) SOCP

$$(P) \quad \begin{array}{ll} \inf & c^T x \\ \text{s.t.} & Ax = b, \\ & x \in Q^{n+1} \end{array}$$

$$(D) \quad \begin{array}{ll} \sup & b^T y \\ \text{s.t.} & c - \sum_i y_i a_i \in Q^{n+1} \end{array}$$

A more commonly known form:

Let $a_i = (u_i, a_{i,1}, \cdots, a_{i,n})$

$c = (v, d_1, \cdots, d_n)$

Then, $c - \sum_i y_i a_i \in Q^{n+1}$ is the same as

$$(v - u^T y, d - \bar{A}^T y) \in Q^{n+1} \iff v - u^T y \geq \| d - \bar{A}^T y \|_2$$

(2) SDP

$$(P) \quad \begin{array}{ll} \inf & C \cdot X \\ \text{s.t.} & A_i \cdot X = b_i \\ & X \in S^n_+ \end{array}$$

$C, A_i \in S^n$, and $A \cdot B = \text{Tr}(AB) = \sum_{i,j} A_{ij} B_{ij}$ is an inner product on S^n.

$$(D) \quad \begin{array}{ll} \sup & b^T y \\ \text{s.t.} & C - \sum_i y_i A_i \in S^n_+ \end{array}$$

$$\Leftrightarrow \begin{bmatrix} v \\ d \end{bmatrix} - \begin{bmatrix} u^T \\ A^T \end{bmatrix} y \in Q^{n+1}.$$

Quick application: Recall

(RLC) $\quad a^T x \le b \quad \forall (a,b) = (a^0, b_0) + \sum_{j=1}^{\ell} \xi_j (a^j, b_j), \quad \xi \in \mathcal{U} \subseteq \mathbb{R}^\ell$

When $\mathcal{U} = \{ y \in \mathbb{R}^\ell : \|y\|_2 \le 1 \}$, we have

$$(\text{RLC}) \Leftrightarrow \left[\sum_{j=1}^{\ell} ((a^j)^T x - b_j)^2 \right]^{1/2} \le b_0 - (a^0)^T x.$$

$$\Leftrightarrow \quad \| [(a^j)^T x - b_j] \|_2 \le b_0 - (a^0)^T x \quad \longrightarrow \text{SOC constraint}$$

* Why single out LP, SOCP, and SDP?
  - "Standard" Convex problems, many solvers available