MS-EXXXX - Optimisation under uncertainty
Lecture 3

Fabricio Oliveira

Department of Mathematics and Systems Analysis
Aalto University, School of Science

August 2, 2023

Outline of this lecture

Introduction

Adjustable robust optimisation

What is robust optimisation

An alternative paradigm for taking uncertainty into account:

- Permeated by the notion of worst-case;
- Control of the degree of conservatism;
- Parallels with chance constraints and risk measures.

In robust optimisation, feasibility is the key concern

- Can be extended to objective function performance requirements;
- May or may not be scenario-based;
- Static v. adaptable: the presence of recourse decisions;
- Exception: distributionally robust optimisation.

Robust optimisation approaches

The key notion in robust optimisation is that of a uncertainty set

- The "region" within which parameter realisation does not turn the solution infeasible;
- Tractability is closely tied to the geometry of such uncertainty sets.

$$\begin{aligned} \min_x .\ & c^\top x \\ \text{s.t.:}\ & Ax \leq b \\ & x \in X \end{aligned}$$

$$\begin{aligned} \min_x .\ & c^\top x \\ \text{s.t.:}\ & A(\eta)x \leq b,\ \forall \eta \in U \subseteq \Xi \\ & x \in X \end{aligned}$$

$$\begin{aligned} \min_x .\ & c^\top x \\ \text{s.t.:}\ & \max_{\eta \in U \subseteq \Xi} A(\eta)x \leq b \\ & x \in X. \end{aligned}$$

Robust counterparts

Let $\tilde{a}_{ij} \in J_i$ be the elements in the matrix $A_{m \times n}$ subject to uncertainty

- Random variables $\tilde{a}_{ij}$ with "central value" a_{ij} and "maximum deviation" $\hat{a}_{ij}$.
- symmetric and limited support $\tilde{a}_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$

Let $\eta_{ij} = \frac{(\tilde{a}_{ij} - a_{ij})}{\hat{a}_{ij}}$. Thus $\eta_{ij} \in [-1, 1]$ and follows the same distribution as $\tilde{a}_{ij}$, but centred in zero and scaled.

Our robust counterpart[1] is the following bilevel problem:

$$\begin{aligned} \min_{x}.\ & c^\top x \\ \text{s.t.:}\ & a_{ij}x_j + \max_{\eta_i \in U_i} \left\{ \sum_{j \in J_i} \eta_{ij}\hat{a}_{ij}x_j \right\} \leq b_i, \forall i \in [m] \\ & x_j \geq 0,\ \forall j \in [n]. \end{aligned} \tag{RC}$$

[1] Assuming, w.l.g., $a_{ij} \geq 0$, $\forall i \in [m], \forall j \in [n]$

Uncertainty set geometries

Box uncertainty set Soyster (1973)

- Maximum protection level;
- All parameters take their worst-possible value
- Simple, but highly conservative

uncertainty set figure

The uncertainty set is

$$U_i = \{\eta_i : ||\eta_i||_1 \leq |J_i|\} \equiv \{\eta_{ij} : |\eta_{ij}| \leq 1, \forall j \in J_i\}$$

The lower-level problem becomes

$$\max_{\eta_i \in U_i} \left\{ \sum_{j \in J_i} \eta_{ij} \hat{a}_{ij} x_j : |\eta_{ij}| \leq 1, \forall j \in J_i \right\} = \sum_{j \in J_i} \hat{a}_{ij} x_j .$$

Uncertainty set geometries

Ellipsoidal uncertainty set Ben-Tal and Nemirovski (1998)

- Softens extreme-case protection;
- Parametrically controlled;
- Leads to smooth sets;
- (MI)QCPs which may be more computationally demanding.

uncertainty set figure

The uncertainty set is

$$U_i = \{\eta_i : ||\eta_i||_2 \leq \Gamma\} \equiv \left\{ \eta_{ij} : \sum_{j \in J_i} \eta_{ij}^2 \leq \Gamma^2, \forall j \in J_i \right\}$$

The lower-level problem becomes

$$\max_{\eta_i \in U_i} \left\{ \sum_{j \in J_i} \eta_{ij} \hat{a}_{ij} x_j : \sum_{j \in J_i} \eta_{ij}^2 \leq \Gamma^2, \forall j \in J_i \right\}.$$

Uncertainty set geometries

Ellipsoidal uncertainty set Ben-Tal and Nemirovski (1998)

Again, this uncertainty set has a closed-form solution:

$$
\begin{aligned}
&\max_{\eta_i \in U_i} \left\{ \sum_{j \in J_i} \eta_{ij} \hat{a}_{ij} x_j : \sum_{j \in J_i} \eta_{ij}^2 \leq \Gamma^2, \forall j \in J_i \right\} \\
=&\max_{\eta_i \in U_i} \left\{ \sqrt{\left(\sum_{j \in J_i} \eta_{ij} \hat{a}_{ij} x_j \right)^2} : \sum_{j \in J_i} \eta_{ij}^2 \leq \Gamma^2, \forall j \in J_i \right\} \\
=&\max_{\eta_i \in U_i} \left\{ \sqrt{\left(\sum_{j \in J_i} \eta_{ij} \right)^2 \left(\sum_{j \in J_i} \hat{a}_{ij} x_j \right)^2} : \sum_{j \in J_i} \eta_{ij}^2 \leq \Gamma^2, \forall j \in J_i \right\} \\
=&\Gamma \sqrt{\sum_{j \in J_i} \hat{a}_{ij}^2 x_j^2}
\end{aligned}
$$

Uncertainty set geometries

Polyhedral uncertainty set Bertsimas and Sim (2004)

- Allows for controlling conservatism;
- Retains problem complexity;
- Budget of uncertainty lacks interpretability.

uncertainty set figure

The uncertainty set is

$$U_i = \{\eta_i : ||\eta_i||_1 \leq \Gamma_i\} \equiv \left\{\eta_{ij} : \sum_{j \in J_i} \eta_{ij} \leq \Gamma_i, \forall j \in J_i\right\}.$$

The lower-level problem becomes

$$\max_{\eta_i \in U_i} \left\{\sum_{j \in J_i} \eta_{ij} \hat{a}_{ij} x_j : \sum_{j \in J_i} \eta_{ij} \leq \Gamma_i,\ 0 \leq \eta_{ij} \leq 1,\ \forall j \in J_i\right\}.$$

Uncertainty set geometries

Polyhedral uncertainty set Bertsimas and Sim (2004)

In this case, the lower-level problem does not admit a closed form but is an LP.

- Strong duality (primal-dual equivalence) is available;
- True for any convex[2] lower-level problem.

$$
\begin{aligned}
\max_{\eta_i \in U_i}.\ & \sum_{j \in J_i} \eta_{ij} \hat{a}_{ij} x_j \\
\text{s.t.:}\ & \sum_{j \in J_i} \eta_{ij} \leq \Gamma_i \ (\pi_i) \\
& 0 \leq \eta_{ij} \leq 1, \ (p_{ij}) \ \forall j \in J_i
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
\min_{\pi_i, p_i}.\ & \Gamma_i \pi_i + \sum_{j \in J_i} p_{ij} \\
\text{s.t.:}\ & \pi_i + p_{ij} \geq \hat{a}_{ij} x_j, \forall j \in J_i \\
& p_{ij} \geq 0, \ \forall j \in J_i \\
& \pi_i \geq 0.
\end{aligned}
$$

[2]Satisfying some constraint qualification.

Robust counterparts

Let us consider a knapsack problem of the form

$$\begin{aligned} \text{min. } & c^\top x \\ \text{s.t.: } & \sum_{j \in [n]} a_j x_j \leq b \\ & 0 \leq x_j \leq 0, \ \forall j \in [n]. \end{aligned}$$

The robust counterparts for the previous uncertainty sets are:

Box

$$\begin{aligned} \text{min. } & c^\top x \\ \text{s.t.: } & \sum_{j \in [n]} (a_j + \hat{a}_{ij}) x_j \leq b \\ & 0 \leq x_j \leq 1, \ \forall j \in [n]. \end{aligned}$$

Ellipsoid

$$\begin{aligned} \text{min. } & c^\top x \\ \text{s.t.: } & \sum_{j \in [n]} a_j x_j + \Gamma \sum_{j \in [n]} (\hat{a} x_j)^2 \leq b \\ & 0 \leq x_j \leq 1, \ \forall j \in [n]. \end{aligned}$$

Robust counterparts

Let us consider a knapsack problem of the form

$$\begin{aligned}\text{min. } & c^\top x \\ \text{s.t.: } & \sum_{j\in[n]} a_j x_j \leq b \\ & 0 \leq x_j \leq 1, \ \forall j \in [n].\end{aligned}$$

The robust counterparts for the previous uncertainty sets are:

Polyhedral

$$\begin{aligned}\text{min. } & c^\top x \\ \text{s.t.: } & \sum_{j\in[n]} a_j x_j + \Gamma\pi + \sum_{j\in[n]} p_j \leq b \\ & \pi + p_j \geq \hat{a}_j x_j, \ \forall j \in [n] \\ & 0 \leq x_j \leq 1, p_j \geq 0, \ \forall j \in [n] \\ & \pi \geq 0.\end{aligned}$$

On constraint violation probabilities

Arguably, Bertsimas & Sim (2004) raised attention to robust optimisation with "the price of robustness".

- The price refers to the optimality traded off for feasibility guarantees;
- Quantifying these trade-offs can be done:
  1. Using theoretical bounds;
  2. Via simulating solution performance.
- In my own experience, theoretical bounds are often loose.

For example, Li et al. (2012) show the probability of violation of constraint $i \in [m]$ to be

$$P^{\text{vio}} = P\left(a_{ij}x_j + \max_{\eta_i \in U_i}\left\{\sum_{j \in J_i} \eta_{ij}\hat{a}_{ij}x_j\right\} > b_i\right) \leq e^{\frac{-\Delta^2}{2|J_i|}},$$

where $\Delta = 1$ for box sets, and Γ for ellipsoidal and polyhedral sets.

Multi-stage robust optimisation

Introducing stages in robust optimisation models can be achieved:

1. Define uncertainty sets via their geometry;
2. Using scenarios.

We focus on 2-stage adjustable robust optimisation (ARO) problems:

$$
\begin{aligned}
\min\ & c^\top x + \max_{\xi \in U \subset \Xi} \min_{y} q^\top y(\xi) \\
\text{s.t.: } & Ax = b,\ x \geq 0 \\
& Tx + Wy(\xi) = h(\xi),\ \forall \xi \in U \subset \Xi \\
& y(\xi) \geq 0,\ \forall \xi \in \Xi.
\end{aligned}
\tag{ARO}
$$

- Only RHS uncertainty: $T(\xi) = T$, $W(\xi) = W$, and $q(\xi) = q\ \forall \xi \in \Xi$;
- Assumption often necessary to eliminate quadratic dependence between ξ and decision variables;
- Not necessary if the uncertainty set is discrete and finite (scenarios)

A side note: mini-max, minimum regret and related

If the uncertainty set is a finite and discrete set of scenarios, we have that

$$\begin{aligned} \min_x.\ & c^\top x + \max_{s \in U} q_s^\top y_s \\ \text{s.t.: } & T_s x + W_s y_s \leq h_s,\ \forall s \in U \\ & x \in X \end{aligned}$$

is a tractable ARO (Mulvey, 1995). Variants include:

Minimax

$$\begin{aligned} \min_x.\ & c^\top x + \theta \\ & \theta \geq q_s^\top y_s,\ \forall s \in U \\ \text{s.t.: } & T_s x + W_s y_s \leq h_s,\ \forall s \in U \\ & x \in X \end{aligned}$$

Min-regret

$$\begin{aligned} \min_x.\ & c^\top x + \theta \\ & \theta \geq q_s^\top y_s - q_s^\top y_s^\star,\ \forall s \in U \\ \text{s.t.: } & T_s x + W_s y_s \leq h_s,\ \forall s \in U \\ & x \in X \end{aligned}$$

where $y_s^\star$ is optimal for $s \in U$.

Affinely adjustable robust optimisation (Ben-Tal et. al 2004)

One idea for modelling adjustability is using affine policies

- Replace $y(\xi)$ with $\alpha + \beta\xi$;
- $h(\xi)$ is assumed affinely dependent on ξ, e.g.: $h(\xi) = h + \hat{h}\xi$.

Then ARO becomes:

$$\begin{aligned}
\min_{x,\alpha,\beta} \ & c^\top x + \theta \\
\text{s.t.: } & \theta \geq q^\top(\alpha + \beta\xi), \ \forall \xi \in U \\
& Ax = b, \ x \geq 0 \\
& Tx + W(\alpha + \beta\xi) \leq h(\xi), \ \forall \xi \in U \\
& y(\xi) \geq 0, \ \forall \xi \in \Xi.
\end{aligned}$$

Similar to the static case, computational tractability can be achieved:

- Requires that U is a box or ellipsoidal set;
- For a practical example, see Ben-Tal et al. (2005)

Adjustable robust optimisation

Another approach consists of looking at the inner problem again as a bilevel problem.

Let us restate our ARO in a simplified notation. For that, let

- $X = \{x \in \mathbb{R}^{n_1} : Ax = b, x \geq 0\}$;
- $Y = \{y \in \mathbb{R}^{n_2} : y \geq 0\}$;
- $\xi \in \left[\underline{\xi}, \overline{\xi}\right]$.

Then we have that ARO is equivalent to

$$\min_{x \in X} c^\top x + \mathcal{Q}(x), \tag{ARO}$$

where

$$\mathcal{Q}(x) = \left\{\max_{\xi \in U} \min_{y \in Y} q^\top y : Tx = (h - \hat{h}\xi) - Wy\right\}.$$

Adjustable robust optimisation & CCG

Let us assume that an oracle is available such that, for a given $\overline{x} \in X$ it evaluates $\mathcal{Q}(x)$ and returns associated $(\overline{\xi}, \overline{y})$, if they exist.

In addition, let us assume that the uncertainty set is finitely representable:

- Scenarios, but an intractable lot of them (e.g., samples, or data)
- Polyhedral set (finite extreme points and rays)

We can employ column-and-constraint generation (CCG) (Zhang and Zhao, 2013) to solve ARO:

Main problem M^k: $\overline{x}^{k+1}$

$$\begin{aligned} \min_{x,y,\theta} \ & c^\top x + \theta \\ & \theta \geq q^\top y_l, \ l \in [k] \\ & x \in X \\ & Tx = h - \hat{h}\overline{\xi}_l - Wy_l, \ l \in [k] \\ & y_l \in Y, \ l \in [k]. \end{aligned}$$

Oracle $\mathcal{Q}(\overline{x}^{k+1})$: $\overline{\xi}^{k+1}$

$$\begin{aligned} \max_{\xi \in U} \min_{y \in Y} \ & q^\top y \\ \text{s.t.: } & T\overline{x}^{k+1} = (h - \hat{h}\xi) - Wy. \end{aligned}$$

Adjustable robust optimisation & CCG

In summary, the CCG method can be stated as

1. **Initialisation.** $LB = -\infty$, $UB = \infty$, $k = 0$.
2. **Solve** the main problem M^k Let $\underline{z}^k = c^\top x^k + \theta^k$, where $(x^k, \theta^k, (y_l^k)_{l=1}^k)$ form the optimal solution of M^k. Make $LB = \underline{z}^k$.
3. **Call the oracle** that solves $\mathcal{Q}(x^k)$. Let $(\overline{\xi}^{k+1}, \overline{y}^{k+1})$ be the optimal solution to $\mathcal{Q}(x^k)$, if it exists. Let $\overline{z}^k = c^\top x^k + \mathcal{Q}(x^k)$. Make $UB = \min\left\{UB, \overline{z}^k\right\}$. If $UB = LB < \epsilon$, **return** x^k.
4. If $\mathcal{Q}(x^k)$ is feasible, create variables y_{k+1} and add them to the main problem, together with the constraints

$$\theta \geq q^\top y_{k+1},\ Tx = h - \hat{h}\overline{\xi}_{k+1} - W y_{k+1},\ y_{k+1} \in Y.$$

to form M^{k+1}. Make $k = k + 1$ and return to Step 2. If $\mathcal{Q}(x^k)$ is not feasible, then only the second constraint needs to be created.

Adjustable robust optimisation & CCG

Practical remarks

Essentially, CCG for ARO is a delayed generation of the min-max formulation

- Can thus be useful when too many scenarios are available;
- Convergence relies on a finiteness argument on the uncertainty set.

CCG can be seen as a primal equivalent to Benders decomposition

- One can use the same column generation approach in the context of the L-Shaped method;
- This can help as a way to transmit "recourse information" to the main problem

Adjustable robust optimisation & CCG

On solving $\mathcal{Q}(x)$

Recall that $\mathcal{Q}(x)$ is of the form

$$\begin{aligned}
\mathcal{Q}(x) = \max_{u}\, & q^\top y \\
\text{s.t.: } & u \in \mathcal{U} \\
& y \in \operatorname*{argmin}_{y} q^\top y \\
& \qquad \text{s.t.: } Tx = h - \hat{h}\xi - Wy \\
& \qquad\qquad y \in Y.
\end{aligned}$$

Which is a bilevel model and can be solved using dedicated methods.

- Most techniques rely on posing optimality conditions of the lower-level problem to yield an equivalent single-level (tractable) problem;
- Thus, lower-level convexity (plus CQ) is often a requirement.

Adjustable robust optimisation & CCG

On solving $\mathcal{Q}(x)$

Example: assume that $Y = \mathbb{R}^{n_2}_+$. We can use strong duality to reformulate the lower-level problem, obtaining

$$\begin{aligned} \mathcal{Q}(x) = &\max_{\xi,\pi} (h - \hat{h}\xi - Tx)^\top \pi \\ \text{s.t.: } &\pi^\top W \leq q^\top \\ &\xi \in U, \end{aligned}$$

which is solvable, if:

1. ξ is integer or has a discrete domain, since $\xi^\top \pi$ can be reformulated exactly (e.g., Rintamaki et al. 2023);
2. if $-(\hat{h}\xi)^\top \pi + (h - Tx)^\top \pi$ is a concave bilinear function in π and ξ;
3. if applying a global solver (e.g., Gurobi's spatial branch-and-bound method) is feasible from a computational standpoint.