Operations research:
Mathematics and algorithmics for solving decision-making problems

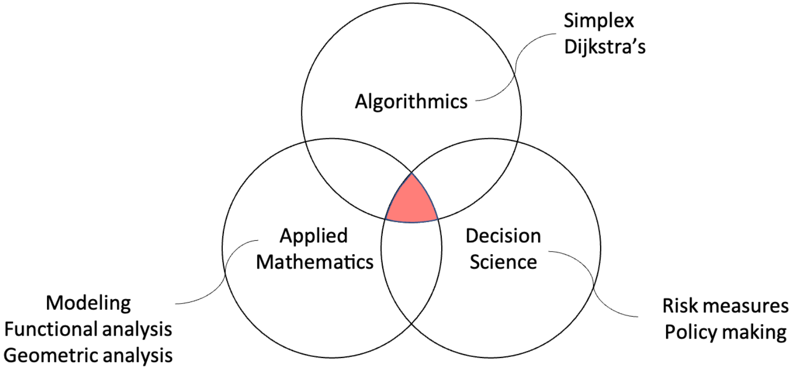
Ayşe N. Arslan

Chargée de recherche
Centre Inria de l'Universite de Bordeaux
Equipe : EDGE

Midis Math
28/11/2023

Operations research (OR)

- Operations research and optimization are at the intersection of multiple disciplines.



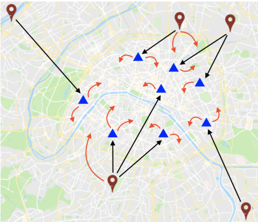
Operations research (OR)

- OR solves optimization problems that decision-makers (managers, politicians, engineers, etc.) encounter.
- Problems are typically expressed in terms of decisions, costs and constraints.
- Tools coming from mathematics, informatics, economics and industrial engineering are often used in their solution.
- The end result is a decision-making tool.

$$\min_{\mathbf{x} \in X} f(\mathbf{x})$$

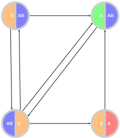
Some examples

- Route optimization
- Planning and scheduling
- Network design (telecommunications, distribution, electricity, etc.)
- Supply chain management



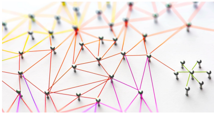
Some projects from our team[1]

Kidney exchange



Maintenance planning

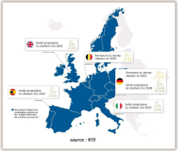
Retail network design



Phytosanitary treatments

Maritime transportation

Energy mix planning



[1] Team EDGE-Centre Inria de l'Universite de Bordeaux

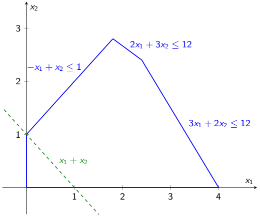
Classical tools of operations research[2]

- Mathematical programming
  - Linear programming (LP/PL)
  - Mixed-integer linear programming (MILP/PLNE)
  - ..
- Graph theory and algorithms
- Constraint programming (CP/PPC)
- Convex analysis
- Approximation algorithms
- Heuristics, metaheuristics
- Queueing theory, simulation, statistics
- ...

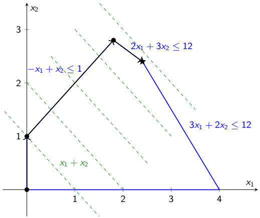
[2] Most topics are covered in the master MAS-ROAD

Mathematical programming: Linear programming

$$\min_{\mathbf{x}\in\mathbb{R}^n_+} \quad \mathbf{c}^\top \mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}$$



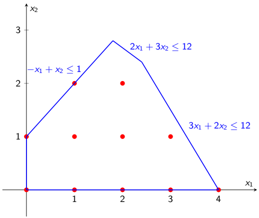
- Well-known algorithms: Simplex, ellipsoid, interior point, etc.
- Solvers: CPLEX, Gurobi, Clp, HiGHS, Excel-Solver etc.



$$z = \mathbf{c}_\mathbf{B}^\top \mathbf{A}_\mathbf{B}^{-1}\mathbf{b} + \left(\mathbf{c}_\mathbf{N} - \mathbf{c}_\mathbf{B}^\top \mathbf{A}_\mathbf{B}^{-1}\mathbf{A}_\mathbf{N}\right)^\top \mathbf{x}_N$$
$$\mathbf{x}_\mathbf{B} = \mathbf{A}_\mathbf{B}^{-1}\mathbf{b} - \mathbf{A}_\mathbf{B}^{-1}\mathbf{A}_\mathbf{N}\mathbf{x}_\mathbf{N}$$

Mathematical programming: Mixed-integer programming

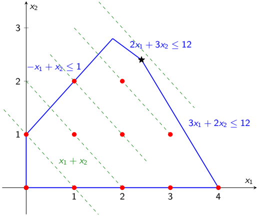
$$\min_{\mathbf{x}\in\mathbb{Z}_+^{n_1}\times\mathbb{R}_+^{n_2}} \quad \mathbf{c}^\top\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}$$

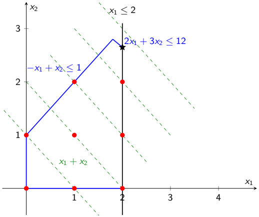


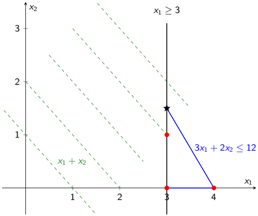
- Integer/binary decisions that cannot be rounded from fractions:
  - Do we open facility i or not?
  - How many trucks do we need to send to client j from facility i?
  - If facility i is not open then it cannot be used to satisfy demand.
- Well-known algorithms: Branch & Bound, Branch & Cut etc.
- Solvers: CPLEX, Gurobi, HiGHS, GLPK etc.

Mathematical programming: Mixed-integer programming

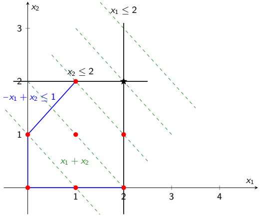
- Branch & Bound: Solve relaxations and successively partition the feasible region

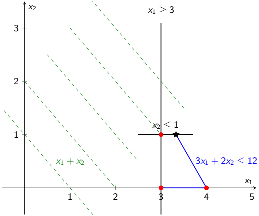


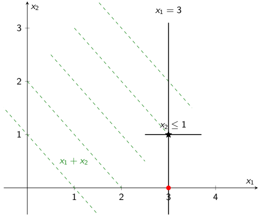


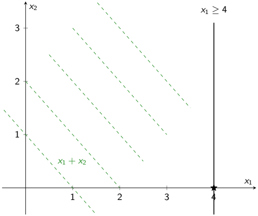


Mathematical programming: Mixed-integer programming









Difficulty of mixed-integer programming problems[3]

- Mixed-integer programming problems are NP-Complete in the general case.
- In practice, considerable progress has been made since the beginning.

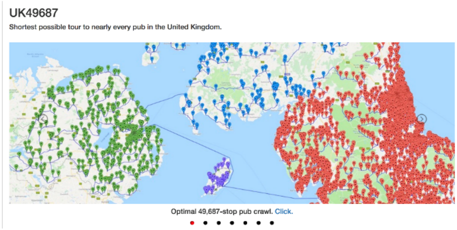
A classical problem: Traveling salesman

- Find the shortest cycle passing through N cities given the pairwise distances.
- In theory finding the best cycle requires testing $N!$ possibilities.
  - For 10 cities (N=10) : < 1 milliseconds
  - For 30 cities (N=30) : 35000 billion years

[3] Source : http://www.math.uwaterloo.ca/tsp/

Difficulty of mixed-integer programming problems[3]

- Mixed-integer programming problems are NP-Complete in the general case.
- In practice, considerable progress has been made since the beginning.

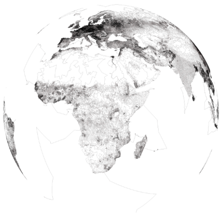


[3] Source : http://www.math.uwaterloo.ca/tsp/

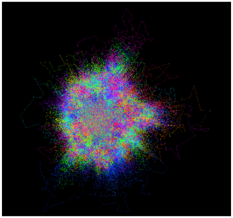
Difficulty of mixed-integer programming problems[3]

- Mixed-integer programming problems are NP-Complete in the general case.
- In practice, considerable progress has been made since the beginning.

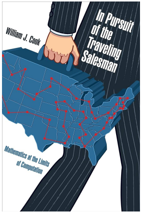
1.9 millions cities to 0.0473% optimality



1 331 906 450 stars to 0.37% optimality

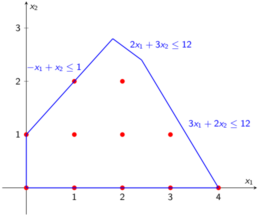


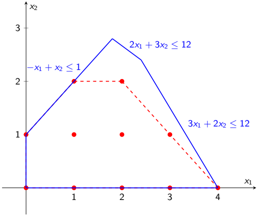
- Powerful heuristics coupled with branch & cut
- Many results dedicated to the problem
- Months of computation in parallel processing



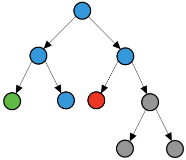
[3] Source : http://www.math.uwaterloo.ca/tsp/

Advanced techniques in mixed-integer programming

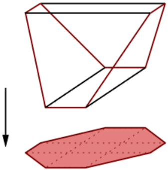




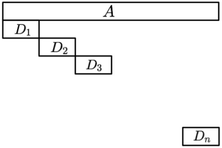
Implicit enumeration



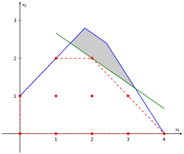
Extended formulations



Decomposition algorithms

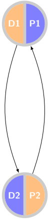


Geometric analysis



What is the Kidney Exchange Problem (KEP)?

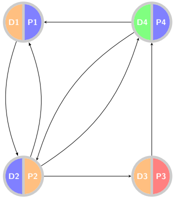
- Chronic kidney disease is a serious health condition threatening the lives of many in our society.
- It is the 11th most common cause of death in the world.
- There are two common treatments for this disease : dialysis and kidney transplant.
- Dialysis is easily available but requires weakly visits to the hospital and degrades the quality of life of the patient.
- Kidney transplants, once successfully performed may give the patients the chance for a healthy life.
- Finding a willing and compatible donor can be very difficult.



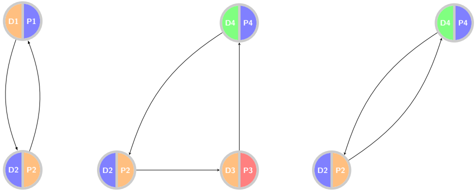
What is the Kidney Exchange Problem (KEP)?

- Given
  - A list of incompatible patient-donor pairs
  - The compatibility information between donors and patients of different pairs
  - A limit K on the size of cycles
- Find a set of exchange cycles such that
  - Each cycle has size $\leq K$
  - Each patient-donor pair is involved in at most one cycle
  - A maximum number of transplants is performed

Instance with 4 pairs, K=3



Possible solutions



This problem is known to be NP-Complete for $K \in [3, \infty)$.

A possible mathematical formulation of the problem

- Data
  - $\mathcal{P}$: set of patient-donor pairs indexed $i = 1, \ldots, |\mathcal{P}|$
  - $(i, j) \in \mathcal{A}$: compatibility information between donors and patients of different pairs
  - K: cycle size limit
- Variables
  - $x_{ij} \in \{0, 1\}$: 1 if a transplant between donor of pair i and patient of pair j is to be performed, 0 otherwise

$$\begin{aligned}
\max \quad & \sum_{(i,j)\in\mathcal{A}} x_{ij} & \\
\text{s.t.} \quad & \sum_{j\in\delta^+(i)} x_{ij} = \sum_{j\in\delta^-(i)} x_{ji} & \forall i \in \mathcal{P} \\
& \sum_{j\in\delta^+(i)} x_{ij} \leq 1 & \forall i \in \mathcal{P} \\
& x \in \{0,1\}^{|\mathcal{A}|} &
\end{aligned}$$

A possible mathematical formulation of the problem

- Data
  - $\mathcal{P}$: set of patient-donor pairs indexed $i = 1, \ldots, |\mathcal{P}|$
  - $(i, j) \in \mathcal{A}$: compatibility information between donors and patients of different pairs
  - K: cycle size limit
- Variables
  - $x_{ij} \in \{0, 1\}$: 1 if a transplant between donor of pair i and patient of pair j is to be performed, 0 otherwise

$$
\begin{aligned}
\max \quad & \sum_{(i,j)\in\mathcal{A}} x_{ij} \\
\text{s.t.} \quad & \sum_{j\in\delta^+(i)} x_{ij} = \sum_{j\in\delta^-(i)} x_{ji} && \forall i \in \mathcal{P} \\
& \sum_{j\in\delta^+(i)} x_{ij} \leq 1 && \forall i \in \mathcal{P} \\
& \color{red}{\sum_{(i,j)\in Path} x_{ij} \leq K - 1} && \color{red}{\forall Path \in AllPaths(K)} \\
& x \in \{0, 1\}^{|\mathcal{A}|}
\end{aligned}
$$

Another possible mathematical formulation

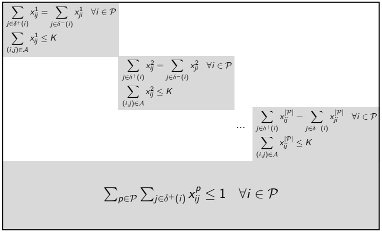
- Data
  - $\mathcal{P}$: set of patient-donor pairs indexed $i = 1, \ldots, |\mathcal{P}|$
  - $(i,j) \in \mathcal{A}$: compatibility information between donors and patients of different pairs
  - K: cycle size limit
- Additional data
  - $\mathcal{C}_K$: set of cycles of size at most K
  - $\mathcal{C}_K(i)$: set of cycles of size at most K containing pair i
  - $|c|$ for $c \in \mathcal{C}_K$: number of pairs involved in cycle c
- Variables
  - $x_c \in \{0,1\}$: 1 if cycle $c \in \mathcal{C}_K$ is chosen as part of the solution, 0 otherwise

$$\begin{aligned} \max \quad & \sum_{c \in \mathcal{C}_K} |c| x_c \\ \text{s.t.} \quad & \sum_{c \in \mathcal{C}_K(i)} x_c \leq 1 \qquad && \forall i \in \mathcal{P} \\ & x \in \{0,1\}^{|\mathcal{C}|} \end{aligned}$$

Yet another possible mathematical formulation

- Data
  - $\mathcal{P}$: set of patient-donor pairs indexed $i = 1, \ldots, |\mathcal{P}|$
  - $(i,j) \in \mathcal{A}$: compatibility information between donors and patients of different pairs
  - K: cycle size limit
- Variables
  - $x_{ij}^p \in \{0,1\}$: 1 if a transplant between donor of pair i and patient of pair j is to be performed in a cycle starting from pair p, 0 otherwise

$$
\begin{aligned}
\max \quad & \sum_{p \in \mathcal{P}} \sum_{(i,j) \in \mathcal{A}} x_{ij}^p \\
\text{s.t.} \quad & \sum_{j \in \delta^+(i)} x_{ij}^p = \sum_{j \in \delta^-(i)} x_{ji}^p && \forall p \in \mathcal{P}, \forall i \in \mathcal{P} \\
& \sum_{(i,j) \in \mathcal{A}} x_{ij}^p \leq K && \forall p \in \mathcal{P} \\
& \sum_{p \in \mathcal{P}} \sum_{j \in \delta^+(i)} x_{ij}^p \leq 1 && \forall i \in \mathcal{P} \\
& x \in \{0,1\}^{|\mathcal{A}| \times |\mathcal{P}|}
\end{aligned}
$$

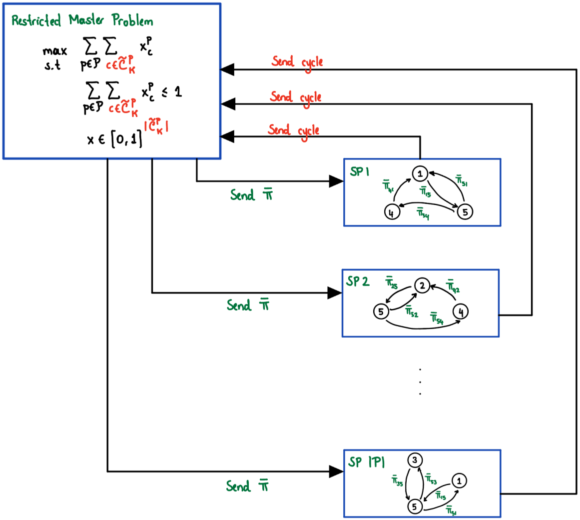


And finally...

- Data
  - $\mathcal{P}$: set of patient-donor pairs indexed $i = 1, \ldots, |\mathcal{P}|$
  - $(i, j) \in \mathcal{A}$: compatibility information between donors and patients of different pairs
  - K: cycle size limit
- Additional data
  - $\mathcal{C}_K^p$: set of cycles of size at most K starting from pair p
  - $\mathcal{C}_K^p(i)$: set of cycles of size at most K starting from pair p containing pair i
  - $|c|$ for $c \in \mathcal{C}_K^p$: number of pairs involved in cycle c
- Variables
  - $x_c^p \in \{0, 1\}$: 1 if cycle c starting from pair p is chosen as part of the solution, 0 otherwise

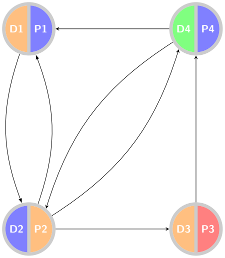
$$
\begin{aligned}
\max \quad & \sum_{p \in \mathcal{P}} \sum_{c \in \mathcal{C}_K^p} |c| x_c^p & \\
\text{s.t.} \quad & \sum_{p \in \mathcal{P}} \sum_{c \in \mathcal{C}_K^p(i)} x_c^p \leq 1 & \forall i \in \mathcal{P} \\
& x \in \{0, 1\}^{|\mathcal{C}_K^p|} & \forall p \in \mathcal{P}
\end{aligned}
$$

Column generation algorithm



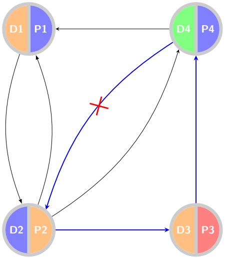
What about uncertainty?

- In reality the compatibility information is only known with certainty after performing costly tests.



What about uncertainty?

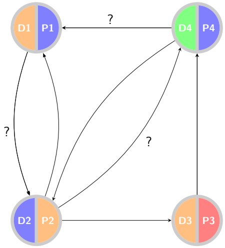
- In reality the compatibility information is only known with certainty after performing costly tests.



- Constructed cycles can break after these tests are performed.

What about uncertainty?

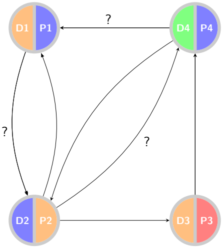
- In reality the compatibility information is only known with certainty after performing costly tests.



- Constructed cycles can break after these tests are performed.
- A better strategy is to test first and then construct a solution with the acquired compatibility information.

What about uncertainty?

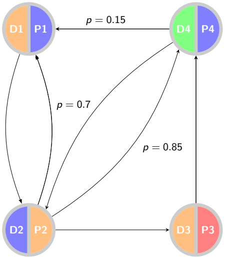
- In reality the compatibility information is only known with certainty after performing costly tests.



- Constructed cycles can break after these tests are performed.
- A better strategy is to test first and then construct a solution with the acquired compatibility information.
- Then a question arises as to what donor-patient pairs to test with a limited budget.

What about uncertainty?

- In reality the compatibility information is only known with certainty after performing costly tests.



- Constructed cycles can break after these tests are performed.
- A better strategy is to test first and then construct a solution with the acquired compatibility information.
- Then a question arises as to what donor-patient pairs to test with a limited budget.

Requires development of optimization under uncertainty approaches.

Conclusions

- OR is an interdisciplinary field that studies decision-making problems from a mathematical and algorithmic perspective.
- It develops tools to help decision makers.
- Significant algorithmic progress has been made in recent years.
- Further research is needed in order to extend classical results to more realistic contexts.

Questions?

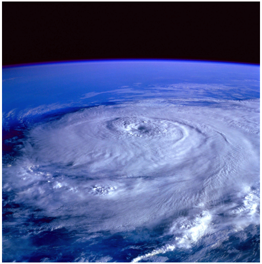
Thank you for your attention!
Any questions?

ayse-nur.arslan@inria.fr
https://www.inria.fr/fr/edge

Appendix:
Optimisation under uncertainty

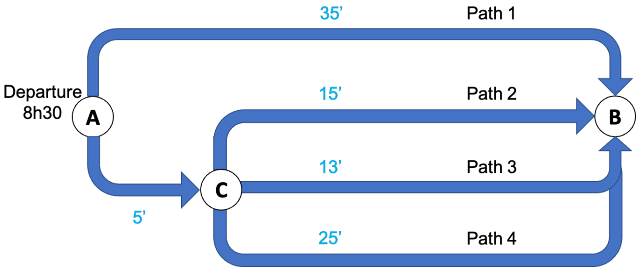
Presence of uncertainty

- New challenges:
  - Renewable energy production.
  - Resilient network design.
  - Healthcare/disaster management.
  - Circular economy.
  - Security and defense.



- Uncertainty:
  - Stochastic nature of systems.
  - Long duration of decision processes.
  - Difficulty of precise measurements.
  - Lack of historical information.
  - Presence of adversarial participants.

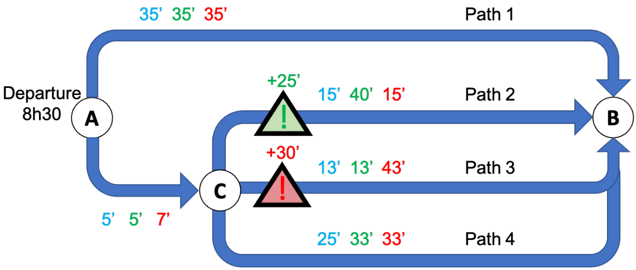
A practical example: Shortest path under uncertainty



Path 1	35′
Path 2	20′
Path 3	18′
Path 4	30′

- What is the shortest path from point A to point B?

A practical example: Shortest path under uncertainty

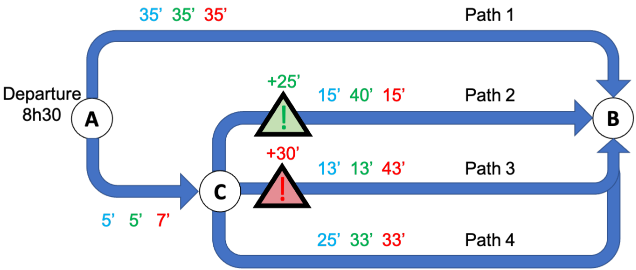


Scenario	Probability
1	0,5
2	0,25
3	0,25

- What is the shortest path from point A to point B?

Our first order of business is to characterize the uncertain data.

A practical example: Shortest path under uncertainty



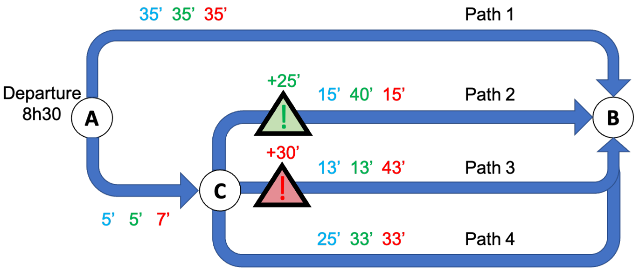
Scenario	Probability
1	0,5
2	0,25
3	0,25

Scenario	1	2	3
Path 1	35'	35'	35'
Path 2	20'	45'	22'
Path 3	18'	18'	50'
Path 4	30'	38'	40'

- What is the shortest path from point A to point B?

Second order of business is to characterize what constitutes a "good" solution.

A practical example: Shortest path under uncertainty



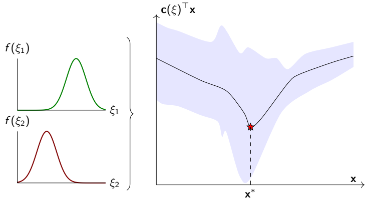
Scenario	Probability
1	0,5
2	0,25
3	0,25

Scenario	1	2	3	Esp	Max
Path 1	35'	35'	35'	35'	35'
Path 2	20'	45'	22'	26,75'	45'
Path 3	18'	18'	50'	26'	50'
Path 4	30'	38'	40'	34,5'	38'

In optimization under uncertainty the notion of a good solution depends on the risk preferences of the decision-maker.

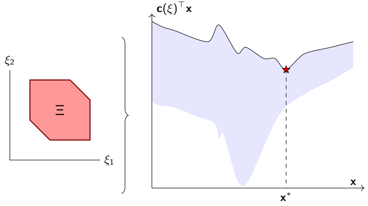
Optimization under uncertainty paradigms

Stochastic optimization



- Distribution $\mathbb{P}$ is known.
- Consequences are observed repeatedly.
- Risk level is low or moderate.
- Example: Distribution network design.

Robust optimization



- Distribution is not known (or no distribution).
- Consequences are observed once.
- Risk level is high.
- Example: Disaster management.

Optimization under uncertainty paradigms

Stochastic optimization

$$\min_{\mathbf{x} \in X} \quad \mathbb{E}^{\mathbb{P}}_{\xi \in \Xi} \left[\mathbf{c}(\xi)^\top \mathbf{x} \right]$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}$$

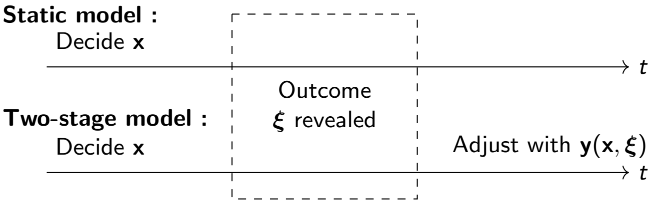
- Distribution $\mathbb{P}$ is known.
- Consequences are observed repeatedly.
- Risk level is low or moderate.
- Example: Distribution network design.

Robust optimization

$$\min_{\mathbf{x} \in X} \quad \max_{\xi \in \Xi} \left[\mathbf{c}(\xi)^\top \mathbf{x} \right]$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}$$

- Distribution is not known (or no distribution).
- Consequences are observed once.
- Risk level is high.
- Example: Disaster management.

Sequential decision-making under uncertainty



In optimization under uncertainty the timing of decisions is important.

The difficulty of solution can increase with the number of decision stages.