Robust optimization modeling

15.C57/15.C571/6.C57/6.C571/IDS.C57: Optimization

Saurabh Amin
Alexandre Jacquillat

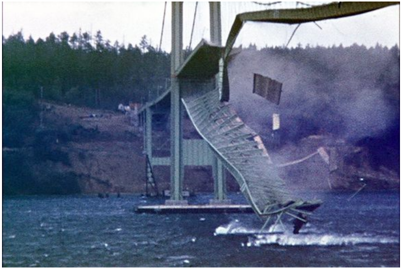
Massachusetts Institute of Technology

Motivation

Real-world examples

Tacoma Narrows Bridge

- Third longest suspension bridge in the world at the time
- Could withstand 120 mph winds, but collapsed in a 42 mph wind due to resonance phenomenon



Scheduling the Panama canal

- Ship scheduling, with extra throughput of 6 large ships/day
- But what happens if some of the vessels break down or some of the ropes tying the ships break?

Data uncertainty in optimization

- A constraint from PILOT4 in the NETLIB library: $\overline{\boldsymbol{a}}^\top \boldsymbol{x} \geq b$

$$
\begin{gathered}
-15.79081 \cdot x_{826} - \mathbf{8.598819} \cdot x_{827} - 1.88789 \cdot x_{828} - 1.362417 \cdot x_{829} \\
-1.526049 \cdot x_{830} - 0.031883 \cdot x_{849} - 28.725555 \cdot x_{850} - 10.792065 \cdot x_{851} \\
-0.19004 \cdot x_{852} - 2.757176 \cdot x_{853} - 12.290832 \cdot x_{854} + 717.562256 \cdot x_{855} \\
-0.057865x \cdot x_{856} - 3.785417 \cdot x_{857} - 78.30661 \cdot x_{858} - 122.163055 \cdot x_{859} \\
-6.46609 \cdot x_{860} - 0.48371 \cdot x_{861} - 0.615264 \cdot x_{862} - 1.353783 \cdot x_{863} \\
-84.644257 \cdot x_{864} - 122.459045 \cdot x_{865} - 43.15593 \cdot x_{866} - 1.712592 \cdot x_{870} \\
-0.401597 \cdot x_{871} + \mathit{1} \cdot x_{880} - 0.946049 \cdot x_{898} - 0.946049 \cdot x_{916} \geq 23.387405
\end{gathered}
$$

→ "Nominal" solution

$$
\begin{aligned}
x^*_{826} &= 255.6112787181108, & x^*_{827} &= 6240.488912232100, \\
x^*_{828} &= 3624.613324098961, & x^*_{829} &= 18.20205065283259, \\
x^*_{849} &= 174397.0389573037, & x^*_{870} &= 14250.00176680900, \\
x^*_{871} &= 25910.00731692178, & x^*_{880} &= 104958.3199274139, \ldots
\end{aligned}
$$

- Data uncertainty
  - Numbers such as *1* are probably certain.
  - Numbers such as **8.598819** are estimated and potentially inaccurate.

→ Will the nominal solution remain optimal? Will it even be feasible?

Impact of data uncertainty

- Worst-case view: What is the largest possible constraint deviation?
  - Suppose the coefficients are 0.1% inaccurate: $|a_i - \overline{a}_i| \leq 0.001|\overline{a}_i|$
  - → The nominal solution can violate the constraint by up to 550%:

$$\min_{\boldsymbol{a}}(\boldsymbol{a}^\top \boldsymbol{x}^* - b) = -128.8 \simeq -5.5b$$

- Average view: What is a "typical" constraint violation?
  - Suppose the coefficients are $a_i = \overline{a}_i + \overline{\varepsilon}_i|\overline{a}_i|$, with $\overline{\varepsilon}_i \sim \mathcal{U}[-0.001, 0.001]$
  - Simulation with 1,000 samples, and constraint violation:

$$\overline{V} = \max\left[\frac{b - \boldsymbol{a}^\top \boldsymbol{x}^*}{|b|}, 0\right]$$

  - → Probability of violated constraint: $\mathbb{P}(\overline{V} > 0) = 50\%$
  - → Average constraint violation of 125%: $\mathbb{E}(\overline{V}) = 1.25$
  - → Worst constraint violation of 450%: $\sum_{i=1}^n a_i x_i^* - b = -104.9 \simeq -4.5b$

- From NETLIB library, with 0.01% uncertainty in model parameters:
  - For $19/90$ problems, some constraints are violated by more than 5%
  - For $13/90$ problems, some constraints are violated by more than 50%

Robust optimization: motivation

- Uncertainty is ubiquitous in decision-making
  - Inexact data (e.g., missing entries, incomplete records)
  - Measurement error (e.g., inaccuracy, rounding, privacy)
  - Estimation error (e.g., demand forecasting, cost/price prediction)
  - Implementation error (e.g., optimization of length, weight, voltage)
- Limitations of probabilistic models in optimization under uncertainty
  - Data: probability distributions are not available in practice
  - Analytical challenges: probability of constraint violation hard to derive
  - Simulation is an approximation and may be computationally expensive
  - Lack of scalability as number of variables and constraints grow

$\rightarrow$ Toward robust optimization

1. Solutions with feasibility guarantees and strong performance even if parameter estimates are wrong due to uncertainty or errors
2. Methodology for optimization under uncertainty that remains tractable in high dimensions

Robust optimization formulations

Robust modeling

Decision **Adversarial deviation** **Adversarial outcomes**

Definition

Uncertainty set $\mathcal{U}$: set of possible value of uncertain parameters that the robust optimization formulation needs to "protect" against

- Robust optimization: worst-case view of uncertainty
  - Parameter values are determined by an adversary
  - The adversary is constrained to select worst-case parameter values within an uncertainty set, defined by the modeler

$\rightarrow$ Modeling: How to define uncertainty sets to represent uncertainty?

$\rightarrow$ Tractability: How to solve the robust optimization problem?

Robust optimization: uncertainty on the constraint matrix

Formulation (Linear optimization)

$$
\begin{aligned}
\min \quad & c_1x_1 + \cdots + c_nx_n \\
\textit{s.t.} \quad & a_{11}x_1 + \cdots + a_{1n}x_n \le b_1 \\
& a_{21}x_1 + \cdots + a_{2n}x_n \le b_2 \\
& \cdots \\
& a_{m1}x_1 + \cdots + a_{mn}x_n \le b_m \\
& x_1, \cdots, x_n \ge 0
\end{aligned}
$$

Formulation (Robust optimization)

$$
\begin{aligned}
\min \quad & c_1x_1 + \cdots + c_nx_n \\
\textit{s.t.} \quad & \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n \le b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n \le b_2 \\ \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \le b_m \end{cases} \\
& \qquad \forall (a_{11}, \cdots, a_{mn}) \in \mathcal{U} \\
& x_1, \cdots, x_n \ge 0
\end{aligned}
$$

$\rightarrow$ Seeking the best possible solution that maintains feasibility for all combination of (uncertain) parameters within the uncertainty set $\mathcal{U}$

$$(a_{11}, \cdots, a_{mn}) \in \mathcal{U}$$

Robust optimization: general uncertainty

- General uncertainty: what if the cost parameters $c_1, \cdots, c_n$ and the resource parameters $b_1, \cdots, b_m$ are also uncertain?
- Consider the linear optimization problem

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

- We can move all parameters to the left-hand size of the constraints:

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & \boldsymbol{c}^\top \boldsymbol{x} - t \leq 0 \\ & \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \leq \boldsymbol{0} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

→ Through appropriate transformations, we can assume that the objective function and the right-hand side are certain, without loss of generality

→ We can focus on robust optimization with uncertainty on the matrix $\boldsymbol{A}$

Robust optimization: general formulation

Formulation (Nominal problem)

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \textit{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

Formulation (Robust problem)

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \textit{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \ \forall \boldsymbol{A} \in \mathcal{U} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

- In practice, constraints are often independent, so we restrict our attention to independent row-wise uncertainty sets

Formulation (Robust optimization with row-wise uncertainty)

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \textit{s.t.} \quad & \boldsymbol{a}_i^\top \boldsymbol{x} \leq b_i, \ \forall \boldsymbol{a}_i \in \mathcal{U}_i, \ \forall i = 1, \cdots, m \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

Constructing uncertainty sets

Example: capacity expansion

- You are planning expansion into 4 new markets with a \$500M budget
  - Each factory earns revenue over the long term, generating an NPV
  - Construction costs are uncertain. From historical data and experience, you may have access to their average and standard deviation

Market	1	2	3	4
Expected cost (\$M)	120	100	180	140
St. Dev. of cost (\$M)	12	10	18	14
Net present value (\$B)	50	40	60	30

→ Deterministic optimization formulation: integer knapsack problem

$$\begin{aligned} \max \quad & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{s.t.} \quad & 120x_1 + 100x_2 + 180x_3 + 140x_4 \leq 500 \\ & x_1, x_2, x_3, x_4 \in \mathbb{Z}_+ \end{aligned}$$

→ Nominal solution: $x_1^* = 4$, $x_2^* = x_3^* = x_4^* = 0$

→ How to build an uncertainty set to protect against deviations in the cost parameters from their expected values?

First attempt: element-wise deviations

- Uncertainty set by letting each parameter deviate from its mean by a factor Γ times the standard deviation (e.g., $\Gamma = 1, 2, 3$)

$$-\Gamma \leq \frac{a_1 - 120}{12} \leq \Gamma, -\Gamma \leq \frac{a_2 - 100}{10} \leq \Gamma, -\Gamma \leq \frac{a_3 - 180}{18} \leq \Gamma, -\Gamma \leq \frac{a_4 - 140}{14} \leq \Gamma$$

$\rightarrow$ Uncertainty set based on element-wise deviations:

$$\mathcal{U} = \left\{ \boldsymbol{a} \in \mathbb{R}^4 : -\Gamma \leq \frac{a_i - \overline{a}_i}{\sigma_i} \leq \Gamma,\ \forall i = 1, \cdots, 4 \right\}$$

$\rightarrow$ Robust optimization formulation:

$$\begin{aligned} \max \quad & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{s.t.} \quad & a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 \leq 500 \text{ for all } \boldsymbol{a} \in \mathcal{U} \\ & x_1, x_2, x_3, x_4 \in \mathbb{Z}, \end{aligned}$$

- This uncertainty set implies that *all* parameters take worst-case values
- Fair criticism: unlikely that all unknowns take worst values together

Second attempt: Central Limit Theorem

Theorem (Central Limit Theorem (CLT))

X_i i.i.d. with mean μ and standard deviation σ. As $n \to \infty$, we have:

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma \cdot \sqrt{n}} \sim \mathcal{N}(0,1).$$

- Intuitively, the sum of random variables tends to be close to their mean

→ Uncertainty set based on the CLT (even for small values of n)

$$\mathcal{U} = \left\{ \boldsymbol{a} \in \mathbb{R}^n : -\Gamma\sqrt{n} \leq \sum_{i=1}^{n} \frac{a_i - \bar{a}_i}{\sigma_i} \leq \Gamma\sqrt{n} \right\}$$

→ Rather than starting from the axioms of probability theory, we start from one of its conclusions: the Central Limit Theorem

- The parameter Γ still controls the degree of robustness:
  - For $\Gamma = 2$, $\mathbb{P}\left[\boldsymbol{a} \in \mathcal{U}\right] \simeq 0.95$ for large values of n
  - For $\Gamma = 3$, $\mathbb{P}\left[\boldsymbol{a} \in \mathcal{U}\right] \simeq 0.997$ for large values of n

Results and insights

Robust capacity expansion

- Robust formulation

$$\begin{aligned} \max \quad & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{s.t.} \quad & a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 \leq 500 \text{ for all } \boldsymbol{a} \in \mathcal{U} \\ & x_1, x_2, x_3, x_4 \in \mathbb{Z}, \end{aligned}$$

- **Nominal:** deterministic formulation: $\mathcal{U}_0 = \{(\overline{a}_1, \overline{a}_2, \overline{a}_3, \overline{a}_4)\}$
  $\rightarrow$ Nominal solution: $x_1^* = 4$, $x_2^* = x_3^* = x_4^* = 0$
- **Robust 1:** element-wise perturbations

$$\mathcal{U}_1 = \left\{ \boldsymbol{a} = (a_1, a_2, a_3, a_4) : -\Gamma \leq \frac{a_i - \overline{a}_i}{\sigma_i} \leq \Gamma \right\}$$

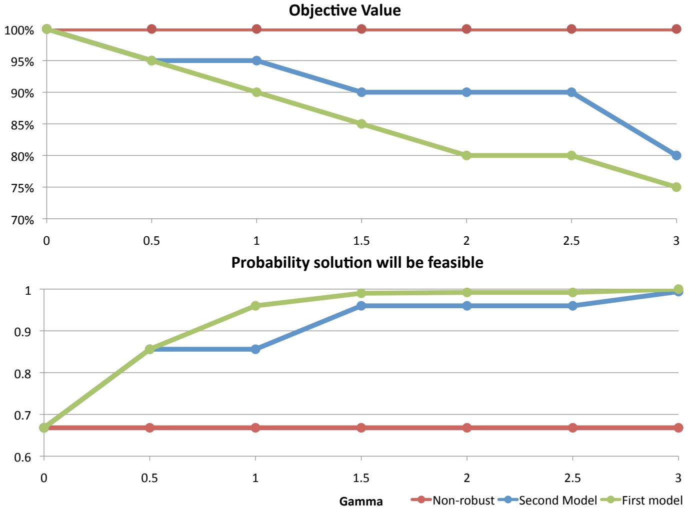
- **Robust 2:** global perturbations, based on the Central Limit Theorem

$$\mathcal{U}_2 = \left\{ \boldsymbol{a} = (a_1, a_2, a_3, a_4) : -\Gamma\sqrt{n} \leq \sum_{i=1}^{n} \frac{a_i - \overline{a}_i}{\sigma_i} \leq \Gamma\sqrt{n} \right\}$$

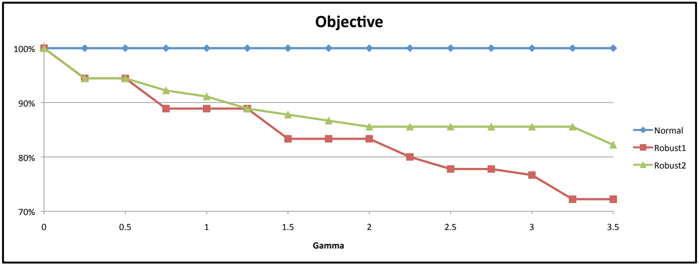
Robust solution

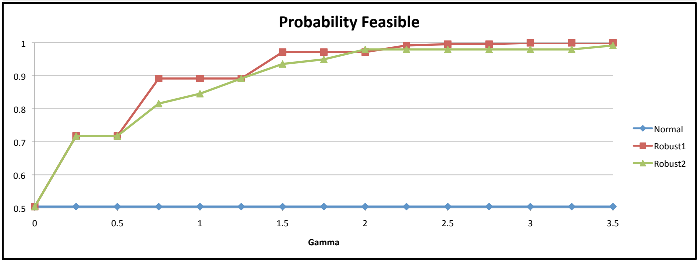
Problem	Γ	Objective	x_1	x_2	x_3	x_4	Unused Capital
Nominal	0	200	4	0	0	0	20
Robust 1	0.5	190	3	1	0	0	40
Robust 1	1	180	2	2	0	0	60
Robust 1	1.5	170	1	3	0	0	80
Robust 1	2	160	0	4	0	0	100
Robust 1	2.5	160	0	4	0	0	100
Robust 1	3	150	3	0	0	0	140
Robust 2	0.5	190	3	1	0	0	40
Robust 2	1	190	3	1	0	0	40
Robust 2	1.5	180	2	2	0	0	60
Robust 2	2	180	2	2	0	0	60
Robust 2	2.5	180	2	2	0	0	60
Robust 2	3	160	2	0	1	0	80

Tradeoff: robustness vs. optimality (4 variables)



Tradeoff: robustness vs. optimality (10 variables)





Observations and insights

- The solution becomes increasingly robust as Γ increases:
  - More unused capital to ensure feasibility in case of higher costs
  - Higher probability that the solution will be feasible
  - Smaller objective value (NPV)
- The price of robustness can be small:
  - For small Γ, the robust solution achieves 90% of the nominal profit while significantly increasing the probability of remaining within budget
- The choice of the uncertainty set matters:
  - The CLT-based uncertainty set restricts the power of the adversary, resulting in a less conservative solution: less unused capital, smaller (but similar) probability of staying within budget, and higher NPV
  - The CLT-based uncertainty set performs even better as n increases, because the CLT becomes a better approximation

  $\rightarrow$ Smaller price of robustness as the number of uncertainties increases
- How to choose the appropriate uncertainty set, e.g., the value of Γ?

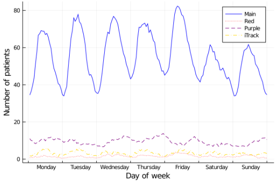
  $\rightarrow$ Akin to hyperparameter tuning in machine learning using validation set

Case study: nurse scheduling

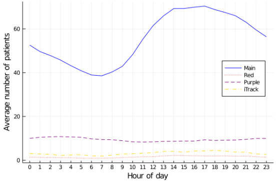
Problem statement

- Nurse staffing: a vital but complex problem for emergency departments
  - Need to avoid understaffing to ensure high quality of healthcare
  - Need to restrict overstaffing to avoid wasting nursing resources
  - Need to ensure satisfaction of nurses based on their preferences
- Challenges to align nurse staffing levels with patient demand
  - Variability in patient demand with time of day and from day to day, creating inconsistencies with (stable) staffing levels
  - Strong uncertainty in patient demand

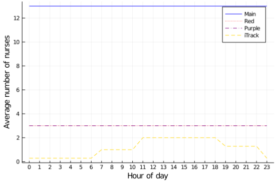
Demand by day



Demand by hour



Current staffing levels



Nurse scheduling optimization

- Nurse scheduling considerations
  - Each nurse belongs to a tier $q \in \mathcal{Q}$ based on experience and qualification
  - Nurses are assigned to work at a certain position $j \in \mathcal{J}$ (e.g., clinical leader, triage nurse, nurse in each pod for each category of patients)
  - Nurses are assigned to work in shift $i \in \mathcal{I}$ on day $d \in \mathcal{D}$

→ Nurse scheduling driven by multi-dimensional assignment variables

$$z_{qjid} = \#\text{ nurses from tier } q \text{ in position } j \text{ during shift } i \text{ on day } d$$

- Scheduling constraints captured by generic representation $\boldsymbol{z} \in \mathcal{Z}$
  - Minimum and maximum staffing levels per position
  - Practical constraints, e.g., eligibility, week-to-week consistency
  - Capacity constraints, e.g., number of available nurses
  - HR constraints, e.g., maximum number of working hours

→ Core nurse scheduling optimization problem:

$$\begin{aligned} \min \quad & \sum_{q \in \mathcal{Q}} \sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{I}} \sum_{d \in \mathcal{D}} z_{qjid} \\ \text{s.t.} \quad & \boldsymbol{z} \in \mathcal{Z} \end{aligned}$$

Handling uncertainty in patient demand

- From historical data, one can estimate target staffing levels:

$$\overline{h}_{jsw} = \#\text{ nurses needed in position } j \text{ at hour } s \text{ on week } w$$

- Uncertainty set to protect against demand variations:
  - ε^1_{js}: maximum forecast error at each hour for each position
  - ε^2_j: maximum forecast error for each week for each position
  - ε^3: maximum forecast error for each week across all positions

$$\mathcal{U} = \left\{ h_{jsw} \geq 0 \;\middle|\; \begin{array}{ll} \left| h_{jsw} - \overline{h}_{jsw} \right| \leq \varepsilon^1_{js}, & \forall j, s, w \\ \left| \sum_{s \in \mathcal{S}_w} h_{jsw} - \sum_{s \in \mathcal{S}_w} \overline{h}_{jsw} \right| \leq \varepsilon^2_j, & \forall j, w \\ \left| \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_w} h_{jsw} - \sum_{j \in \mathcal{J}} \sum_{s \in \mathcal{S}_w} \overline{h}_{jsw} \right| \leq \varepsilon^3, & \forall w \end{array} \right\}$$

Robust optimization formulation

- Multi-objective formulation, with weights μ_1, μ_2
  - Minimizing number of nurse-shifts scheduled
  - Minimizing insufficiency from target nurse-to-patient ratio (npr), using σ_{id}^{sw} to map day d and shift i into week w and hour s
  - Minimizing deviation in staffing levels from current schedule

→ Full robust optimization formulation

$$\min \quad \sum_{q\in\mathcal{Q}}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}}\sum_{d\in\mathcal{D}} z_{qjid} + \mu_1 \cdot npr + \mu_2 \cdot \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}}\sum_{d\in\mathcal{D}} \Delta z_{jid}$$

$$\text{s.t.} \quad npr \geq \sum_{w\in\mathcal{W}}\sum_{s\in\mathcal{S}_w}\sum_{j\in\mathcal{J}} \lambda_w \left(h_{jsw} - \sum_{q\in\mathcal{Q}}\sum_{i\in\mathcal{I}}\sum_{d\in\mathcal{D}} \sigma_{id}^{sw} z_{qjid} \right), \ \forall \boldsymbol{h} \in \mathcal{U}$$

$$\Delta z_{jid} \geq \left| z_{jid}^{cur} - \sum_{q\in\mathcal{Q}} z_{qjid} \right|, \ \forall j \in \mathcal{J},\ i \in \mathcal{I},\ d \in \mathcal{D}$$

$$\boldsymbol{z} \in \mathcal{Z}$$

The edge of robustness

- The robust solution provides significant benefits against benchmarks
  - Current solution: 56 nurse-shifts per day, insufficiency of 0.17 shifts
  - Perfect-information oracle: 11% cost reduction, virtually no insufficiency
  - Deterministic solution: low cost but higher insufficiency
  - Robust solution: 28% reduction in insufficiency with competitive cost (2% higher than deterministic solution, 3% higher than oracle)

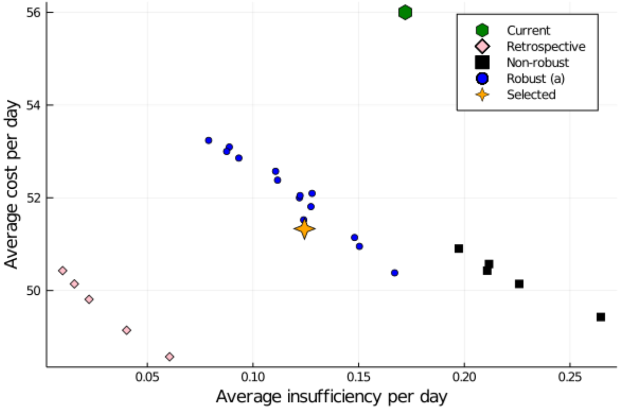
→ Strong impact of robust optimization
  - Significant benefits in terms of nurse-to-patient insufficiency
  - Small price of robustness in terms of staffing costs

Schedule	Cost	Reduction	Insufficiency	Reduction
Current	56	—	0.17	—
Oracle	49.81	11.05%	0.02	86.96%
Deterministic	50.42	9.95%	0.21	-22.55%
Robust	51.33	8.33%	0.12	27.70%

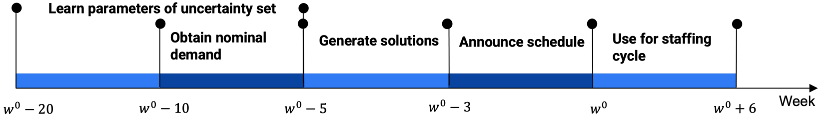
Trade-off between cost and insufficiency

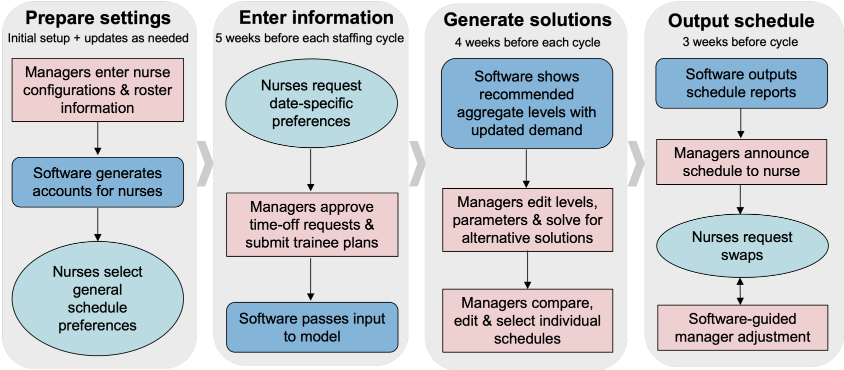
- By varying the hyperparameter μ_1, we identify the trade-off between daily cost and nurse-to-patient insufficiency across all solutions

$\rightarrow$ Selection of a solution that achieves a Pareto improvement over the current solution: less insufficiency and lower cost



Deployment and implementation





Conclusion

Summary

Takeaway

Robust optimization is a tractable methodology to find decisions that are immune to data uncertainty, estimation errors, and implementation errors.

Takeaway

Robust modeling protects against adversarial perturbations in uncertain parameter values, within an uncertainty set.

Takeaway

The uncertainty set is built from the conclusions of probability theory, not its axioms, and it therefore does not rely on distributional assumptions.

Takeaway

The price of robustness is often small: by sacrificing a bit of optimality, the robust solution can ensure feasibility for a vast range of uncertainty.