15.094J: Robust Modeling, Optimization, Computation

Lectures 1: Why RO? Probability Theory and its Limitations

Outline

1. Administration
2. Motivation
   - Data Uncertainty
   - Robust antenna design
   - Probability theory in 20th Century
   - Philosophy
   - Where Probability Distributions Exist?
   - Tractability of Stochastic Models
   - Conclusions
3. Goals
4. The Class Lecture by Lecture

Objectives Today

- Administration
- Why RO?
- What are the new primitives of uncertainty?
- What do robust solutions look like and what they offer?

Administration

- **Time**: Wednesdays, 9:00-12:00.
- **Place**: E51-372.
- **Professor**: Dimitris Bertsimas, E40-111, tel. 617-253-4223.
- **Office hours**: Wednesdays 12-1 or by appointment, e-mail: dbertsim@mit.edu. http://web.mit.edu/dbertsim/www/
- **TAs**: Jack Dunn, E40, e-mail: jackdunn@mit.edu email
- **Office hours:** Tue-Thur: 1-2pm .
- **Recitation**: Friday 11-12, E51-145.

Administration

- **Text:** Research papers and class notes in https://stellar.mit.edu.
- **Recitations:** Background on optimization and probability, software for RO, computational aspects, examples and applications.
- **Course Requirements:** 30% problem sets, 30% midterm exam, and 40% final team project.
- **Background required:** Mathematical maturity. Knowledge of optimization and probability we will develop in recitations.

Why RO?

- Data Uncertainty
- Implementatiion Error
- Computational Limitations of Probability Theory in higher dimensions

Motivation: Data Uncertainty

- Tacoma Bridge
- Panama Canal
- A constraint from PILOT4 in the NETLIB library:

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

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

Motivation: Data Uncertainty

Optimal solution

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

Suppose the coefficients are only 0.1% inaccurate. Will above solution still be feasible?

Motivation: Data Uncertainty

Optimal solution

$$
\begin{aligned}
x^*_{826} &= 255.6112787181108, & x^*_{827} &= 6240.488912232100,\\
x^*_{828} &= 3624.613324098961, & x^*_{829} &= 18.2020506528325,\\
x^*_{849} &= 174397.0389573037, & x^*_{870} &= 14250.00176680900,\\
x^*_{871} &= 25910.00731692178, & x^*_{880} &= 104958.3199274139, \ldots
\end{aligned}
$$

Suppose the coefficients are only 0.1% inaccurate. Will above solution still be feasible?

Simulating with 1000 samples:

$$P(\text{constraint infeasible}) = 50\%,$$

$$\text{worst infeasibility} \ : \sum_{i=1}^{n} a_i x_i^* - b = -104.9, \quad \text{as opposed to } \geq 0.$$

Motivation: Data Uncertainty

Optimal solution

$$\begin{array}{ll} x^*_{826} = 255.6112787181108, & x^*_{827} = 6240.488912232100, \\ x^*_{828} = 3624.613324098961, & x^*_{829} = 18.2020505283259, \\ x^*_{849} = 174397.0389573037, & x^*_{870} = 14250.00176680900, \\ x^*_{871} = 25910.00731692178, & x^*_{880} = 104958.3199274139, \ldots \end{array}$$

Suppose the coefficients are only 0.1% inaccurate. Will above solution still be feasible?

Simulating with 1000 samples:

$$P(\text{constraint infeasible}) = 50\%,$$

$$\text{worst infeasibility} \;:\; \sum_{i=1}^{n} a_i x_i^* - b = -104.9, \quad \text{as opposed to } \geq 0.$$

From NETLIB Library:

- For 27/90 problems, solutions were *infeasible* for very small perturbations
- For 17/90 problems, solutions were infeasible by at least 50%.

Implementation Errors: Robust antenna design

- In a 2-D plane, antennae radiation patterns (i.e., the *radiation diagram*) vary with the incidence angle θ.
- Using a weighted array of linear antenna components, we can control the shape of the resulting diagram. Specifically, the diagram for a linear array of N elements spaced apart at distance d has the form

$$Z(\theta) = \sum_{l=0}^{N-1} w_l \exp\left(j\left(\frac{2\pi d}{\lambda}\right) l \sin\theta\right),$$

  where w_l are the weights of the components (our design variables) and λ is the wavelength of the radiation.
- The weights w_l are generally complex-valued (i.e., we control both magnitude and phase).

Robust antenna design

- Often we wish to keep $Z(\theta)$ large in some angle range around a target angle θ_t and small outside this range.
- In other words, with Δ as the angular "beam-width," keep $Z(\theta)$ close to one (normalized) for $\theta \in [\theta_t - \Delta, \theta_t + \Delta]$, and close to zero for θ outside this range.
- Typical design problem: minimize the *uniform norm*, i.e., minimize the maximum deviation from the desired target pattern over all angles of interest. The *nominal design problem* then is

$$\text{minimize} \qquad \max_{\theta \in \Theta} |Z(\theta) - \delta(\theta)|,$$

  where $\Theta = \{\theta \mid |\theta - \theta_t| \geq \Delta\} \cup \{\theta_t\}$ and $\delta(\theta) = 1$ if $\theta = \theta_t$ and 0 otherwise.
- With complex design weights, this problem may be solved efficiently.

Implementation error

- In practice, however, we cannot implement the design weights with exact precision. One simple way to model this is to assume the actual weights $\tilde{w}_l$ are of the form

$$\tilde{w}_l = (1 + \xi_l) w_l,$$

where w_l are the desired weights, and ξ are zero-mean random variables capturing the imperfect implementation precision.

- In general, these uncontrollable perturbations can have an adverse effect on the performance of the nominal design. The *robust design problem* wishes to protect against all such perturbations within some "reasonable" uncertainty set.

Robust formulation

- The specific uncertainty set we use is

$$P_\gamma = \{\boldsymbol{\xi} \mid \|\boldsymbol{\xi}\| \leq \gamma\},$$

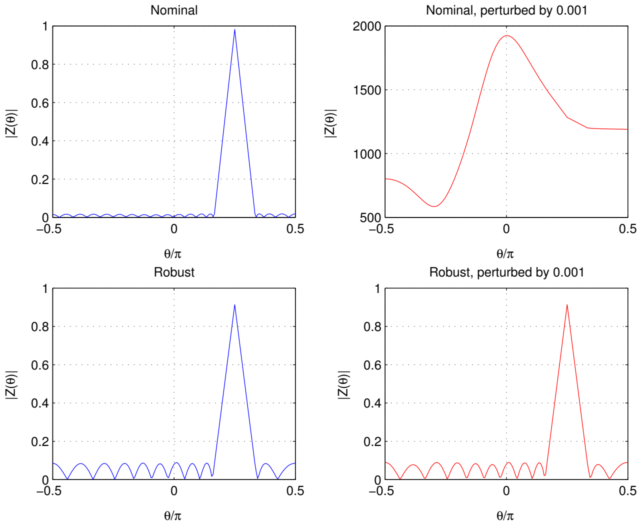
parameterized by some protection level γ.

- The *robust design problem at level* γ, then, is

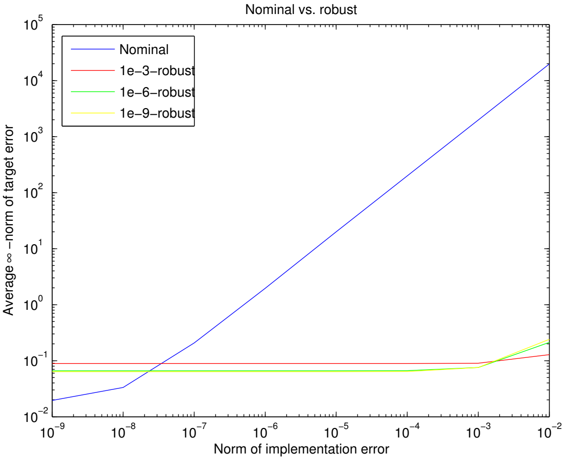
$$\text{minimize} \quad \max_{\theta \in \Theta, \boldsymbol{\xi} \in P_\gamma} |\tilde{Z}(\theta, \boldsymbol{\xi}) - \delta(\theta)|,$$

where $\tilde{Z}(\theta, \boldsymbol{\xi})$ is the antenna diagram perturbed by implementation error $\boldsymbol{\xi}$.

$N = 20$, $d = 1.2\lambda$, $\theta_t = \pi/4$, and $\Delta = \pi/12$, $\gamma = 0.001$



Robust vs Nominal



Summary

- Even extremely small perturbations can have adverse effects on the performance of an antenna design which ignores these implementation errors.
- Robust solutions may be calculated efficiently. These solutions offer vastly superior protection in the face of random errors and do not perform much worse than the nominal designs in the absence of implementation errors.

Major applications of applied probability in the 20th century

Successes	Challenges
Queueing Theory: Erlang (1909) Modeling queues with exponential distributions in steady-state	Transient analysis, analysis of queueing networks with general distributions
Information Theory: Shannon (1948) Characterization of the capacity of a single user channel	Network Information theory
Market Design: Myerson (1981) Characterization of the optimal auction for single item and multiple buyers	Multiple items for buyers with budget constraints, and correlated valuations
Options Pricing: Black and Scholes (1973) Pricing options in low dimensions in complete markets	Pricing options in high dimensions in incomplete markets
Stochastic Optimization: Bellman (1957) Dynamic Programming, successful in low dimensions	Extending computation to high dimensions
Low Dimensions	*High Dimensions*

Philosophy / Motivation

- Performance analysis given that primitives are probability distributions does not extend in a computationally tractable way in *high dimensions*.
- Combining probability theory and optimization often leads to the *"Curse of dimensionality"*.
- Contrast with linear, convex, even discrete *optimization*.
- What was *Kolmogorov*'s intention?
- What was *Dantzig*'s intention?

Past Voices

"All epistemological value of the theory of probability is based on this: that large scale random phenomena in their collective action create strict, non random regularity"– Gnedenko and Kolmogorov (1954)

"If a queue has an arrival process which cannot be well modeled by a Poisson process or one of its near relatives, it is likely to be difficult to fit any simple model, still less to analyze it effectively. So why do we insist on regarding the arrival times as random variables, quantities about which we can make sensible probabilistic statements? Would it not be better to accept that the arrivals form an irregular sequence, and carry out our calculations without positing a joint probability distribution over which that sequence can be averaged? " – J.F.C. Kingman (2009)

Where Probability Distributions Exist ?

- What is available in practice is **data**, not probability distributions.
- Probability distributions are not known in practice; they exist in *our imagination*.
- Thus, modeling with probability distributions is a *choice*. Is it the right one?
- But even if they are available, what happens if the future distribution is not the same as in the past?
- Examples: Hurricane Katrina and the Chermobyl nuclear disaster.

Example: A Capacity Expansion Problem

You are considering entering 4 new markets. You could build multiple factories (or none) in each market, but you have a finite budget of \$500 M. The cost of building a factory differs by market (see below).

If you build a factory, you'll earn revenue over the long term, whose NPV is also represented below.

Market	1	2	3	4
Cost (\$M / factory)	120	100	180	140
NPV Future Revenue (\$B)	50	40	60	30

How many factories should you build in each market?

Example: A Capacity Expansion Problem (continued)

You might solve the following optimization problem:

$$\begin{array}{rl} \text{maximize} & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{subject to} & 120x_1 + 100x_2 + 180x_3 + 140x_4 \leq 500 \\ & x_i \text{ integer.} \end{array}$$

Solution: $x_1^* = 4$, $x_2^* = x_3^* = x_4^* = 0$.

Example: A Capacity Expansion Problem (continued)

In reality, the upfront costs and the projected revenues may not be known exactly. Considering just the costs, suppose they are distributed according to a distribution $f(x)$.

Project	1	2	3	4
Expected Cost	120	100	180	140
St Dev of Cost	12	10	18	14

- How would you compute the probability that the above solution won't break the budget?
- What if the costs are correlated?
- How would you find the best solution which is feasible at least 95 % of the time?

Tractability of Stochastic Modeling

- Even for simple distributions and small problems, calculating the probability of violation of a constraint is often difficult.
- Simulation is only an approximation and may be very computationally expensive.
- Problems get worse as number of variables and constraints grow.
- Optimization in these environment is even harder.
- Another example: Options pricing in hIgh dimensions.

Conclusions; Why RO?

- In my experience in the real world, *robustness* is often more important than *optimality*.
- *Motivation I:* Create solutions that are immune to implementation errors and data uncertainty.
- *Motivation II:* Develop a theory of performance analysis and optimization under uncertainty via optimization that is tractable in high dimensions.
- A remark on *Tractability:* We do not mean polynomial solvability. Rather the ability to solve problems of the size that we care in applications and within computational times appropriate for the application.

Objectives of the Class:

- Propose an alternative to stochastic *modeling* via RO to model uncertain phenomena.
- Develop RO as a tractable methodology for solving *optimization* problems under uncertainty.
- Cover a large number of *applications*.
- Expose you to large scale *computation* for RO.

Lectures

Lecture	Time	Topic	Readings
1	W, 2/08	Probability theory and its limitations	[1]
2	W, 2/08	The new primitives: Uncertainty sets	[1]
3	W, 2/15	Robust linear optimization I	[11]
4	W, 2/15	Robust linear optimization II	[29, 14]
5	W, 2/22	Robust mixed integer optimization	[30, 31]
6	W, 2/22	Robust convex optimization	[10, 32]
7	W, 3/01	Data driven robust optimization	[22]
8	W, 3/01	From data to decisions	[25]
9	W, 3/08	Adaptive optimization I	[9, 23, 8]
10	W, 3/08	Adaptive optimization II	[24, 15, 18, 17]
11	W, 3/15	Distributionally robust optimization	[33]
12	W, 3/15	Distributionally adaptive optimization	[33]
13	W, 3/22	Power of robust policies in adaptive optimization	[21, 20]

Lectures

Lecture	Time	Topic	Readin
14	W, 3/22	Power of affine policies in adaptive optimization	[19, 2
	W, 3/29	Spring break	
15	W, 4/05	RO, risk preferences and utilities	[13, 2
16	W, 4/05	RO in energy	[26]
17	W, 4/12	Robust portfolios	[36, 2
18	W, 4/12	Robust options pricing	[3]
19	W, 4/19	Robust steady-state queueing theory	[6]
20	W, 4/19	Robust transient queueing theory and supply chains	[7, 34
	W, 4/26	Exam	
21	W, 5/03	Robust mechanism design	[2]
22	W, 5/03	RO in statistics	[35, 1
23	W, 5/10	RO in control theory	[12]
24	W, 5/10	Robust information theory	[4, 5]
25	W, 5/17	Project presentations	
26	W, 5/17	Project presentations	

C. Bandi and D. Bertsimas.
Tractable stochastic analysis in high dimensions via robust optimization.
Mathematical Programming, Series B, 134(1):23–70, 2012.

C. Bandi and D. Bertsimas.
Optimal design for multi-item auctions: A robust optimization approach.
Mathematics of Operations Research, 39(4):1012–1038, 2014.

C. Bandi and D. Bertsimas.
Robust option pricing - An ϵ-arbitrage approach.
European Journal of Operational Research, 239(3):842–853, 2014.

C. Bandi and D. Bertsimas.
Channel coding via robust optimization, Part I: the single channel case.
Submitted to *IEEE Transactions on Information Theory*, 2016.

C. Bandi and D. Bertsimas.
Channel coding via robust optimization, Part II: the multiple channel case.
Submitted to *IEEE Transactions on Information Theory*, 2016.

C. Bandi, D. Bertsimas, and N. Youssef.
Robust queueing theory.
Operations Research, 63(3):676 – 700, 2015.

C. Bandi, D. Bertsimas, and N. Youssef.
Robust transient queueing theory.
Submitted to *Queueing Systems*, 2015.

A. Ben-Tal, B. Golany, E. Guslitzer, and A. Nemirovski.
Adjustable robust solutions of uncertain linear programs.
Mathematical Programming, Series B, 99(2):351–376, 2004.

A. Ben-Tal, A. Goryashko, A. Nemirovski, and J.P. Vial.
Retailer-supplier flexible commitments contracts: A robust optimization approach.
Manufacturing and Service Operations Management, 7(3):248–271, 2005.

A. Ben-Tal and A. Nemirovski.
Robust convex optimization.
Mathematics of Operations Research, 23(4):769–805, 1998.

A. Ben-Tal and A. Nemirovski.
Robust solutions of uncertain linear programs.
Operations Research Letters, 25(1):1–13, 1999.

D. Bertsimas and D. Brown.
Constrained stochastic LQC: A tractable approach.

IEEE Transactions on Automatic Control, 52:1826–1841, 2007.

D. Bertsimas and D. Brown.
Constructing uncertainty sets for robust linear optimization.
Operations Research, 57(6):1483–1495, 2009.

D. Bertsimas, D. Brown, and C. Caramanis.
Theory and applications of robust optimization.
SIAM Review, 53(3):464–501, 2011.

D. Bertsimas and C. Caramanis.
Finite adaptability in linear optimization.
IEEE Transations on Automatic Control, 55(12):2751–2766, 2010.

D. Bertsimas and M. Copenhaver.
Characterization of the equivalence of robustification and regularization in linear, median, and matrix regression.
submitted to *European Journal of Operations Research*, 2016.

D. Bertsimas and I. Dunning.
Multistage robust mixed integer optimization with adaptive partitions.
submitted to *Operations Research*, 2014.

D. Bertsimas and A. Georghiou.

Design of near optimal decision rules in multistage adaptive mixed-integer optimization.
Operations Research, 63 (3):610–627, 2015.

D. Bertsimas and V. Goyal.
On the power and limitations of affine policies in two-stage adaptive optimization.
Mathematical Programming, Series B, 134(2):491–531, 2011.

D. Bertsimas and V. Goyal.
On the approximability of adjustable robust convex optimization under uncertainty.
Mathematical Methods of Operations Research, 77(3):323–343, 2013.

D. Bertsimas, V. Goyal, and X.A. Sun.
A geometric characterization of the power of finite adaptability in multi-stage stochastic and adaptive optimization.
Mathematics of Operations Research, 36(1):24–54, 2011.

D. Bertsimas, V. Gupta, and N. Kallus.
Data-driven robust optimization.
Submitted to *Operations Research*, 2013.

D. Bertsimas, D. Iancu, and P. Parrilo.
Optimality of affine policies in multi-stage robust optimization.
Mathematics of Operations Research, 35(2):363–394, 2010.

D. Bertsimas, D. Iancu, and P. Parrilo.
A hierarchy of near-optimal policies for multi-stage adaptive optimization.
IEEE Transactions on Automatic Control, 56(12):2809–2824, 2011.

D. Bertsimas and N. Kallus.
From predictive to prescriptive analytics.
submitted to *Management Science*, 2015.

D. Bertsimas, E. Litvinov, X.A. Sun, J. Zhao, and T. Zheng.
Adaptive robust optimization for the security constrained unit commitment problem.
IEEE Transactions on Power Systems, 28(1):52–63, 2013.

D. Bertsimas and A. O'Hair.
Learning preferences under noise and loss aversion: An optimization approach.
Operations Research, 61(5):1190–1199, 2013.

D. Bertsimas and D. Pachamanova.

Robust multiperiod portfolio management in the presence of transaction costs.
Computers and Operations Research, 35(1):3–17, 2008.

D. Bertsimas, D. Pachamanova, and M. Sim.
Robust linear optimization under general norms.
Operations Research Letters, 32(6):510–516, 2004.

D. Bertsimas and M. Sim.
Robust discrete optimization and network flows.
Mathematical Programming, Series B, 98(1-3):49–71, 2003.

D. Bertsimas and M. Sim.
The price of robustness.
Operations Research, 52(1):35–53, 2004.

D. Bertsimas and M. Sim.
Tractable approximations to robust conic optimization problems.
Mathematical Programming, Series B, 107(1-2):5–36, 2006.

D. Bertsimas, M. Sim, and M Zhang.
Distributionally adaptive optimization.
2015.

Submitted to *Management Science*.

D. Bertsimas and N. Youssef.
Stochastic optimization in supply chain networks: Averaging robust solutions.
2015.
Submitted to *Management Science*.

L. El-Ghaoui and H. Lebret.
Robust solutions to least-square problems to uncertain data matrices.
SIAM Journal on Matrix Analysis and Applications, 18(4):1035–1064, 1997.

D. Goldfarb and G. Iyengar.
Robust portfolio selection problems.
Mathematics of Operations Research, 28(1):1–38, 2003.