Robust Optimization Seminar

Berk Öztürk 03/05/2024

Introductory Remarks

- About me:
 - Optimizer / Operations Research Analyst at MITRE
 - What I do: optimization, data science, simulation and modeling.
 - PhD, MIT AeroAstro, 2022

"Global and Robust Optimization for Engineering Design" Advisor: Prof. Dimitris Bertsimas. TAed Robust Optimization in Spring 2021.

• Please feel free to ask questions during the talk. I will be asking you questions too! [Paraphrasing mercilessly from lecture notes of MIT's 15.094 RO course, taught Spring 2021 by Dimitris Bertsimas and Dick den Hertog (and TAed by me).]

What to expect from today

- Theory of robust optimization (RO), which
 - turns stochastic optimization problems into deterministic ones,
 - is general and practical for a range of decision problems,
 - Provides guarantees of constraint satisfaction under uncertainty.
- Some demonstrations of the need for RO, and derivations of key concepts.
- Suggestions about how you can use RO in your own work.
- A real-time demonstration of a facility location problem under uncertainty.

Parametric uncertainty is ubiquitous in decision/design problems.

Uncertain parameters

Uncertainty arises from:

- Measurement errors,
 - Blood pressure, temperature ...
- Estimation/prediction errors,
 - Demand, truncation error ...
- Implementation errors,
 - Voltages, engineering tolerances ...







^ "Flaw of Averages", Sam Savage, 2012.

Here's a scenario* to avoid...

 $\bar{a}^{T}x = -15.79081x_{826} - 8.598819x_{827} - 1.88789x_{828} - 1.362417x_{829} - 1.526049x_{830} \\ -0.031883x_{849} - 28.725555x_{850} - 10.792065x_{851} - 0.19004x_{852} - 2.757176x_{853} \\ -12.290832x_{854} + 717.562256x_{855} - 0.057865x_{856} - 3.785417x_{857} - 78.30661x_{858} \\ -122.163055x_{859} - 6.46609x_{860} - 0.48371x_{861} - 0.615264x_{862} - 1.353783x_{863} \\ -84.644257x_{864} - 122.459045x_{865} - 43.15593x_{866} - 1.712592x_{870} - 0.401597x_{871} \\ +x_{880} - 0.96049x_{898} - 0.946049x_{916} \\ > b = 23.387405$

Suppose: accuracy is 0.1%:

$$(*) \qquad \left|a_i^{true} - \bar{a}_i\right| \leq 0.001 \left|\bar{a}_i\right|.$$

Worst case: the constraint can be violated by as much as 450%:

$$\min_{a^{true} \text{satisfies }(*)} \left(a^{true}\right)^T ar{x} - b = -128.8 pprox -4.5b.$$

Optimization under uncertainty (OuU) is critical to protect against adverse outcomes!

*From 15.094 lecture notes, 2021.

My original motivation: Legacy aerospace design methods do not adequately consider the risk-performance tradeoff.



Why don't we tackle the following problem directly?

 $egin{array}{lll} \min & f_0(x) \ {
m s.t.} & f_i(x,u) \leq 0, \ orall u \in \mathcal{U}, \ i=1,\ldots,n \end{array}$

Note: there is no uncertainty in the objective function, without loss of generality (wlog).

Answer: infinitely-many constraints can be generated with the infinite number of possible realizations of u.

One method to approximate the problem is through stochastic optimization.



- Problem: Even low-D problems with nice distributions cause computational trouble. **The Curse of Dimensionality.**
 - Jargon: Propagating distributions through constraints via convolutions and high-dimensional quadrature is computationally expensive.

RO is an alternative for optimization under uncertainty that is tractable and <u>deterministic</u>.

$$egin{array}{lll} \min & f_0(x) \ ext{s.t.} & \max_{u\in\mathcal{U}} f_i(x,u) \leq 0, \ i=1,\ldots,n \end{array}$$



RO makes sure all constraints are feasible for all parameter outcomes from an uncertainty set, while minimizing the worstcase objective.

Important notes and intuitions

- Robust optimization problem is robustified *constraint-wise*:
 - Each constraint protects against **all** possible outcomes, allowing different uncertain outcomes for each constraint.

$$egin{array}{lll} \min & f_0(x) \ ext{s.t.} & \max_{u\in\mathcal{U}} f_i(x,u) \leq 0, \; i=1,\ldots,n \end{array}$$

- Thus, depending on the structure of the uncertainty, RO can be conservative, or not!
- RO makes mild assumptions about what the uncertainty set can be.
- RO is deterministic. I will demonstrate in a second!

Robust (Linear) Optimization Theory

Constraint under uncertainty

$$(a+Pz)^T x \le b, \ \forall z \in \mathcal{Z},$$

Uncertain coefficients

Decision variables

can be transformed via the max,

$$\max_{z \in \mathcal{Z}} (a + Pz)^T x \le b.$$

z is *adversarial* to x, so cannot solve in one shot except through reformulation.

Say our uncertainty comes from a box, i.e. bounded by the inf-norm hypercube... can we find a *robust counterpart?*

$$\max_{z:||z||_{\infty} \leq \rho} (\bar{a} + Pz)^{T} x \leq b,$$

$$\bar{a}^{T} x + \max_{z:||z||_{\infty} \leq \rho} (P^{T} x)^{T} z \leq b,$$
Math identity: Separating
$$\bar{a}^{T} x + \max_{z:||z_{i}| \leq \rho} \sum_{i} (P^{T} x)_{i} z_{i} \leq b,$$
Math identity: Rearranging
$$p = \infty$$

$$\bar{a}^{T} x + \rho \sum_{i} |(P^{T} x)_{i}| \leq b,$$
Optimization: In this case intuition.
$$\bar{a}^{T} x + \rho ||P^{T} x||_{1} \leq b.$$

Robust counterpart does not contain $z \in Z!$ ¹⁴

Practical RLO questions (1/3) $(a+Pz)^T x \le b, \forall \{z: ||z||_{\infty} \le \rho\}$ became $a^T x + \rho ||P^T x||_1 \le b.$

Questions:

- 1. The robust counterpart is what kind of optimization problem?
- 2. What is the complexity of this optimization problem? (Hint: can you estimate the number of constraints? Assume $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$.)

Practical RLO questions (2/3)

 $(a + Pz)^T x \leq b, \forall \{z : ||z||_{\infty} \leq \rho\}$ became $a^T x + \rho ||P^T x||_1 \leq b$.

Questions:

- 1. The robust counterpart is what kind of optimization problem? A linear optimization problem!
- 2. What is the complexity of this optimization problem? (Hint: can you estimate the number of constraints? Assume $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$.) m extra variables and 2m linear constraints are required, on top of the n original variables and constraint.

Practical RLO questions (3/3)

 $(a+Pz)^T x \leq b, \forall z: ||z||_{\infty} \leq \rho$ became $a^T x + \rho ||P^T x||_1 \leq b$.

Let's explicitly show the m extra variables and 2m linear constraints that make up the robust counterpart:

$$a^{T}x + \rho \sum_{i \in |m|} \delta_{i} \leq b,$$

$$\delta_{i} \geq P^{T}x, \forall i \in |m|,$$

$$\delta_{i} \geq -P^{T}x, \forall i \in |m|,$$

$$\delta \in \mathbb{R}^{m}.$$

Takeaway: Added constraints can substantially affect the tractability of the robust counterpart!

Let's try another. How about polyhedral uncertainty?

Constraint with uncertain z in a polyhedron:

$$(a+Pz)^T x \le b, \forall z : Dz \le d.$$
(1)

Take the worst-case,

$$a^T x + \max_{z: Dz \le d} (P^T x)^T z \le b.$$

$$\tag{2}$$

Use the dual!

$$\max_{z: Dz \le d} (P^T x)^T z = \min_{D^T y = P^T x, y \ge 0} d^T y.$$
(3)

Replace max with min:

$$a^T x + \min_{D^T y = P^T x, y \ge 0} d^T y \le b.$$
 (4)

Rearrange:

$$a^{T}x + d^{T}y \le b, D^{T}y = P^{T}x, y \ge 0.$$
 (5)

18

Robust counterparts summary

- Started with infinite-dimensional problem under uncertainty.
- Took the max of the left-hand-side with respect to uncertain parameters.
- Did some math magic to end with deterministic optimization, devoid of uncertain parameters!

Note: <u>You do not have to derive these yourself</u>, unless your uncertainty lies in an exotic set. Many robust counterparts exist in the literature.

Cautions:

- Explicit robust counterparts may or may not exist (but there are workarounds).
- If they exist, it's not obvious that they are tractable (but very often they are).

Some examples of robust counterparts for linear constraints

Unc. set Ζ **Robust Counterpart** Tractability $\bar{a}^T x + \rho \| P^T x \|_1 \le b$ $\|z\|_{\infty} \leq \rho$ Box LO (convex quadratic $\bar{a}^T x + \rho \| P^T x \|_2 \le b$ CQO Ellipsoidal $\|\mathbf{z}\|_2 \le \rho$ optimization) $\begin{cases} \bar{a}^T x + d^T y \le b \\ D^T y = P^T x \\ y \ge 0 \end{cases}$ (linear optimization) Polyhedral Dz < dLO $\begin{aligned} \|z\|_{\infty} &\leq \rho \\ \|z\|_{1} &\leq \Gamma \end{aligned}$ $\|\bar{a}^T x + \rho \|y\|_1 + \Gamma \|P^T x - y\|_{\infty} \le b$ Budget LO My personal favorite. Linear

 $(\bar{a}+Pz)^T x \leq b, \quad \forall z \in Z.$

20

and tractable.

New research allows for finding robust counterparts for almost any combination constraint and set.

- For convex constraints and sets, it's called Reformulation-Perspectification.
- If you are comfortable with mathematics of convexity, please check out the paper:
 - Bertsimas, Dimitris, et al. "Robust convex optimization: A new perspective that unifies and extends." *Mathematical Programming* 200.2 (2023): 877-918.
- For uncertainty sets with discrete partitions, see below:
 - Bertsimas, Dimitris, and Iain Dunning. "Multistage robust mixedinteger optimization with adaptive partitions." Operations Research 64.4 (2016): 980-998.

Sometimes, we don't even know we are using robustness, when we are!

Anyone use LASSO for linear regression (LR)? What's the benefit of LASSO over unregularized linear regression?

$$\min_{\beta} ||\mathbf{y} - \mathbf{X}\beta||_q + \rho ||\beta||_r \quad \text{vs.} \quad \min_{\beta} ||\mathbf{y} - \mathbf{X}\beta||_q$$

LASSO is <u>robust</u> regression (<u>not sparse</u>). (1/2)

Let's derive it!

$$\min_{\beta} \max_{\Delta \mathbf{X} \in \mathcal{U}} ||\mathbf{y} - (\mathbf{X} + \Delta \mathbf{X})\beta||_2.$$
(1)

Let's assume that \mathcal{U} is feature-wise uncorrelated:

$$\mathcal{U} := \left\{ (\delta_1, \dots, \delta_m) : ||\delta_i||_2 \le c_i, \ \forall i \in |m| \right\}.$$
(2)

Separate, with a relaxation:

$$\min_{\beta \in (\delta_1, \dots, \delta_m) : ||\delta_i||_2 \le c_i} \max_{||\mathbf{y} - (\mathbf{X} + (\delta_1, \dots, \delta_m))\beta||_2 \le min_{\beta} \left\{ ||\mathbf{y} - \mathbf{X}\beta||_2 + \max_{(\delta_1, \dots, \delta_m) : ||\delta_i||_2 \le c_i} \sum_{i=1}^m ||\beta_i \delta_i||_2 \right\}$$
(3)

*H. Xu, C. Caramanis, and S. Mannor. Robust regression and Lasso. http://arxiv.org/abs/0811.1790v1, 2008. Also submitted to NeurIPS.

LASSO is <u>robust</u> regression (<u>not sparse</u>). (2/2)

Recall our upper bound on the original robust optimization problem:

$$\min_{\beta} \left\{ ||\mathbf{y} - \mathbf{X}\beta||_2 + \max_{(\delta_1, \dots, \delta_m): ||\delta_i||_2 \le c_i} \sum_{i=1}^m ||\beta_i \delta_i||_2 \right\}$$
(1)

Let's add an additional restriction on the δ_i 's:

$$\delta_i := -c_i \operatorname{sign}(\beta_i) u, \text{ where } u = \begin{cases} \frac{\mathbf{y} - \mathbf{X}\beta}{||\mathbf{y} - \mathbf{X}\beta||_2}, & \text{if } \mathbf{y} \neq \mathbf{X}\beta \\ \text{any vector with unit } 12 - \text{norm otherwise} \end{cases}$$
(2)

This is a type of induced norm that still obeys $||\delta_i||_2 \leq c_i$. It also reduces to

$$||\beta_i \delta_i||_2 = ||-\beta_i c_i \operatorname{sign}(\beta_i) u||_2 = c_i |\beta_i|.$$
(3)

There is a little more derivation in the paper^{*}, but LASSO is *equivalent* to the above robust optimization problem under the induced norm:

$$\min_{\beta} ||\mathbf{y} - \mathbf{X}\beta||_2 + \rho ||\beta||_1 \leftrightarrow \min_{\beta} ||\mathbf{y} - \mathbf{X}\beta||_2 + \sum_{i=1}^m c_i |\beta_i|, \text{ when } c_i = \rho, \forall i \in [m]!$$
(4)

*H. Xu, C. Caramanis, and S. Mannor. Robust regression and Lasso. http://arxiv.org/abs/0811.1790v1, 2008. Also submitted to NeurIPS.



How can you use robust optimization in your work?

Option 1 (medium-hard difficulty): Optimize directly using the robust counterpart.

- Steps
 - Write your optimization problem with no uncertainty.
 - Determine your uncertainty set, ideally using some data or model.
 - Look up the robust counterpart online (or derive it 😁).
 - Rewrite the constraints with uncertainty using the robust counterpart (and additional auxiliary variables and constraints).
 - Solve.
- (+) Fast, deterministic, one-shot solutions.
- (-) Can still be computationally difficult. Lots of work up front to write robust counterparts.

Option 2 (easy-medium difficulty): Use adversarial robust optimization.

- Steps:
 - Solve the nominal problem, with no uncertainty
 - Repeat:
 - For each constraint:
 - Solve the inner maximization problem for worst-case parameters u from the uncertainty set, for fixed decisions x.
 - If constraints are violated, use those parameters to add cuts (i.e. perturb the original constraints and add them to the existing constraints).
 - Terminate if no cuts are added!
 - Re-solve the problem for x with all accumulated cuts.
- (+) Can be applied to any mixed-integer convex optimization problem, without taking the robust counterpart, under any uncertainty set.
- (-) Computationally expensive. Check out this paper for a comparison: Bertsimas, Dimitris, Iain Dunning, and Miles Lubin. "Reformulation versus cutting-planes for robust optimization: A computational study." Computational Management Science 13 (2016): 195-217.

Facility Location Demo

(can find on my GitHub, <u>ROdemos/homeworks/HW4 at master · 1ozturkbe/ROdemos · GitHub</u>)

Problem setup (1/4)

- N = 10 possible facility locations, with supply capacity limits.
- M = 50 customers, with demand of 50.
- Question we are answering: How many facilities should we construct, to serve customer demand while minimizing cost? (A MILO!)
 - Cost = facility fixed costs + transport costs (distance * flow).
- Customers may be served by multiple facilities.
- There is budget uncertainty on demand:

$$\begin{cases} ||d||_{\infty} \le \rho = 1, \\ ||d||_{1} \le \Gamma = 5. \end{cases}$$

Problem setup (2/4)

$$\min_{\boldsymbol{y}(\cdot),\boldsymbol{x}} \quad \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} y_{ij} + \sum_{i=1}^{n} f_{i} x_{i}$$

$$s.t. \quad \sum_{i=1}^{n} y_{ij} \ge d_{j}, \quad \forall d \in \mathcal{D}, \forall j \in [m],$$

$$\sum_{j=1}^{m} y_{ij} \le s_{i} x_{i}, \quad \forall i \in [n],$$

$$y_{ij} \ge 0, \quad \forall i \in [n], \forall j \in [m],$$

$$\boldsymbol{x} \in \{0, 1\}^{n}.$$

Objective with variable and fixed costs

Customer-wise supply and demand constraint

Facility-wise capacity constraint

Transport variables Facility location variables

Problem setup (3/4)

$$\begin{split} \min_{\boldsymbol{y}(\cdot),\boldsymbol{x}} & \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} y_{ij} + \sum_{i=1}^{n} f_{i} x_{i} \\ s.t. & \sum_{i=1}^{n} y_{ij} \geq d_{j}, \quad \forall d \in \mathcal{D}, \forall j \in [m], \\ & \sum_{j=1}^{m} y_{ij} \leq s_{i} x_{i}, \quad \forall i \in [n], \\ & y_{ij} \geq 0, \quad \forall i \in [n], \; \forall j \in [m], \\ & \boldsymbol{x} \in \{0,1\}^{n}. \end{split}$$

Question: Notice anything that could be problematic with this formulation?

Uncertainty is spread over many constraints, thus very conservative. (This will be partially mitigated by adaptivity, so stay tuned!)

$$\begin{array}{l} \begin{array}{l} \begin{array}{l} \displaystyle \operatorname{Problem \ setup \ }(4/4) \\ \displaystyle \min_{\boldsymbol{y}(\cdot),\boldsymbol{x}} \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{j=1}^{n}$$

1 1

 $x \in \{0,1\}^n$.

ve with a convex uncertainty set. But we will assume we NEED to satisfy demand exactly.

33

Addressed via four methods.

- Nominal (no uncertainty)
- Robust (static transportation decision rules)
- Adaptive (affine transportation decision rules w.r.t uncertainty)
 - We fix facility locations *x* before demand is realized, then make an affine transportation decision based on that realized demand *d*.

$$\{y = u + Vd, \ d_j = \bar{d}_j + (Pz)_j, \ \forall \ ||z||_{\infty} \le \rho, ||z||_1 \le \Gamma\}.$$

• Adversarial adaptive (affine transportation decision rules w.r.t uncertainty, via cuts)

Quick intro to adaptive robustness

$$\max_{d \in D} A(d)x + By(d) \le b.$$
(1)

A is an uncertain matrix, x are the first-stage decisions, and y(d) are the decision rules, which we will assume are linear for our application: y(d) = u + Vd. Assuming budget uncertainty, taking the robust counterpart has the same process!

$$\max_{d \in D} (a + Pd)^T x + B^T (u + Vd) \le b.$$

$$\tag{2}$$

$$a^{T}x + B^{T}u + \max_{d \in D} d^{T}P^{T}x + d^{T}BV^{T} \le b.$$
 (3)

If d is in the budget set, i.e. $||d||_{\infty} \leq \rho$, and $||d||_1 \leq \Gamma$, we can simply use the robust counterpart table to rewrite as

$$a^{T}x + B^{T}u + \rho||\xi||_{\infty} + \Gamma||\xi||_{1} \le b, \ \xi = P^{T}x + BV^{T}.$$
(4)

Adaptive RO has high computational cost (especially depending on solver).



Solution time: Adaptive takes 40s with Gurobi, adversarial 2 hours or so.

With GLPK, 40 minutes adaptive, adversarial >24 hours. julia> include("nominal.jl") GLPK Simplex Optimizer 5.0 60 rows, 510 columns, 1010 non-zeros INTEGER OPTIMAL SOLUTION FOUND Facility cost: 21.50916629989814 Transportation cost: 12.55731068395753 0.522297 seconds (397.70 k allocations: 33.663 MiB, 32.59% compilation time)

Scales of problems

julia> @time include("robust.jl") GLPK Simplex Optimizer 5.0 10060 rows, 5560 columns, 23560 non-zeros INTEGER OPTIMAL SOLUTION FOUND Facility cost: 34.35670583607979 Transportation cost: 16.2223369008492 plt = nothing 3.054248 seconds (1.15 M allocations: 78.755 MiB, 0.54% gc time, 10.28% compilation time)

julia> include("adaptive.jl") GLPK Simplex Optimizer 5.0 112761 rows, 82172 columns, 455032 non-zeros INTEGER OPTIMAL SOLUTION FOUND 2308.518484 seconds (3.48 M allocations: 196.539 MiB, 0.00% gc time, 0.00% compilation time) Facility cost: 21.427802592678976 Transportation cost: 13.355109872484778

The price of robustness can be low!



We only increased our cost by 7.5%, to protect against a huge amount demand uncertainty (more precisely up to 14% of total across the 50 customers)!

Engineering Design under Uncertainty via Robust Optimization (from PhD thesis work)

Application: aircraft design that captures important multidisciplinary tradeoffs.

- Unmanned, gas-powered aircraft
- Without uncertainty: 176 variables and 154 constraints
- Monolithic: optimizes aircraft and flight trajectory concurrently through disciplined signomial programming (nonconvex) form

Wing

- Structure
- Fuel volume
- Profile drag
- Stall constraint

Fuselage

- Fuel and payload
- Profile drag

Engine

- Data-based sizing
- Lapse rate
- BSFC fits
- T/O and TOC constraints

First, we determine uncertain parameters and variances, based on engineering intuition.

Parameters	Description	Value	% Uncert. (3σ)
e	span efficiency	0.92	3
μ	air viscosity (SL)	$1.78 imes 10^{-5} \text{ kg/(ms)}$	4
ρ	air density (SL)	1.23 kg/m^3	5
$C_{L,\max}$	stall lift coefficient	1.6	5
k	fuselage form factor	1.17	10
$C_{f,\mathrm{ref}}$	reference fuselage skin friction factor	0.455	10
$\rho_{\rm p}$	payload density	1.5 kg/m^3	10
$N_{ m ult}$	ultimate load factor	3.3	15
V_{\min}	takeoff speed	$35 \mathrm{m/s}$	20
$W_{ m p}$	payload weight	3000 N	20
$W_{\rm coeff, strc}$	wing structural weight coefficient	$2 imes 10^{-5} \ \mathrm{1/m}$	20
$W_{\rm coeff, surf}$	wing surface weight coefficient	60 N/m^2	20

Table 3.1: Parameters and uncertainties (increasing order)

The uncertainty is defined by box and ellipsoidal sets.



Figure 3-5: Γ defines the overall size of norm uncertainty sets, while 3σ defines the relative size of the set in each uncertain parameter.

Primary result: RO mitigates probability of constraint violation under uncertain outcomes,



Free variable	Description	Units	No Uncert.	Margins	Box	Elliptical
L/D	mean lift-to-drag ratio	-	33.6	23.6	25.1	27.7
AR	aspect ratio	-	24.6	13.3	13.0	16.3
Re	Reynolds number	-	1.54×10^6	$2.65 imes10^6$	$3.03 imes 10^6$	$250 imes 10^6$
S	wing planform area	m^2	13.6	32.8	32.0	28.1
V	mean flight velocity	m/s	41.6	37.3	38.9	38.4
$T_{ m flight}$	time of flight	hr	20.1	22.4	21.4	21.7
$W_{ m w}$	wing weight	Ν	2830	4760	4800	4480
$W_{ m w,strc}$	wing structural weight	Ν	2010	4760	2670	2620
$W_{\rm w,surf}$	wing skin weight	Ν	820	2170	2120	1860
$W_{ m fuse}$	fuselage weight	N	250	314	288	279
$V_{ m f,avail}$	total fuel volume	m^3	0.0759	0.146	0.154	0.136
$V_{\rm f,fuse}$	fuselage fuel volume	m^3	0.0394	0	0	0.0159
$V_{\rm f,wing}$	wing fuel volume	m ³	0.0365	0.167	0.154	0.120
Objective metric	Description	Units	No Uncert.	Margins	Box	Elliptical
Objective	total fuel weight	Ν	608	1170	1240	1090
E[Objective]	expected total fuel weight	Ν	572	964	976	856
σ [Objective]	std. dev. of fuel weight	Ν	9	32	32	29
P[failure]	probability of failure	%	94	0	0	0

Table 2: SP Aircraft Optimization Results, for $\Gamma = 1$

*100 MC simulations over 3σ truncated Gaussians

References / resources

ROdemos https://github.com/1ozturkbe/ROdemos

This repository contains a number of practical and tractable examples of robust optimization (RO), applied to problems as diverse as experimental design and multicommodity network flows. They are written in Julia, using the JuMP modeling environment. The models accept a variety of optimizers compatible with JuMP, but Gurobi is the default, available with a free academic license.

Good papers on this topic

- Bertsimas, Dimitris, Iain Dunning, and Miles Lubin. "Reformulation versus cutting-planes for robust optimization: A computational study." Computational Management Science 13 (2016): 195-217.
- Bertsimas, D., & Sim, M. (2004). The Price of Robustness. Operations Research, 52(1), 35-53. <u>https://doi.org/10.1287/opre.1030.0065</u>
- Bertsimas, D., Brown, D. B., & Caramanis, C. (2011). Theory and Applications of Robust Optimization. *SIAM Review*, *53*(3), 464–501. https://doi.org/10.1515/9781400831050
- Bertsimas, D., Hertog, D. den, & Pauphilet, J. (2019). Probabilistic guarantees in Robust Optimization. September, 1–27.
- Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*. https://doi.org/10.1016/S1470-2045(03)01230-0
- Bertsimas, D., Brown, D. B., & Brown, D. B. (2009). Constructing Uncertainty Sets for Robust Linear Optimization. September 2019. https://doi.org/10.1287/opre.1080.0646
- Bertsimas, Dimitris, et al. "Robust convex optimization: A new perspective that unifies and extends." Mathematical Programming 200.2 (2023): 877-918.
- H. Xu, C. Caramanis, and S. Mannor. Robust regression and Lasso. http://arxiv.org/abs/0811.1790v1, 2008. Also submitted to NeurIPS.

Mediocre papers on this topic

 Öztürk, B., & Saab, A. (2021). Optimal Aircraft Design Decisions Under Uncertainty Using Robust Signomial Programming. AIAA Journal, 59(5), 1–43. https://doi.org/10.2514/1.j058724