Optimization Models

EECS 127 / EECS 227AT

Laurent El Ghaoui

EECS department
UC Berkeley

Fall 2018

LECTURE 13

Robust Optimization Models

The future is uncertain . . . but this uncertainty is at the very heart of human creativity.

Ilya Prigogine

Outline

Curse of uncertainty

"Nominal" optimization problem:

$$\min_x \; f_0(x) \;:\; f_i(x) \le 0, \quad i = 1, \ldots, m.$$

In practice, problem data is uncertain:

- *Estimation* errors affect problem parameters.
- *Implementation* errors affect the decision taken.

Uncertainties often lead to highly unstable solutions, or much degraded realized performance.

These problems are compounded in problems with multiple decision periods.

Example

In this example arising in antenna array design, the problem is to approximate a "target" function (in blue) with a linear combination of given "profiles" (functions not shown).

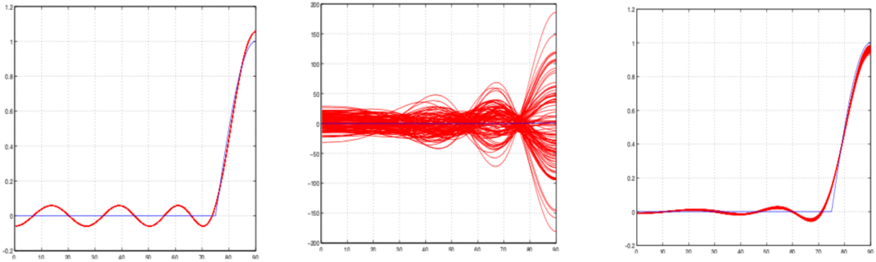


Figure: Antenna design: nominal, perturbed nominal, robust

The nominal solution, when implemented with a .01% relative error, gives a very bad result. A robust approach sacrifices a bit of performance, but completely removes this high sensitivity issue.

Robust counterpart

"Nominal" optimization problem:

$$\min_x \; f_0(x) \;:\; f_i(x) \le 0, \quad i = 1, \ldots, m.$$

Robust counterpart:

$$\min_x \; \max_{u \in \mathcal{U}} \; f_0(x, u) \;:\; \forall\, u \in \mathcal{U}, \quad f_i(x, u) \le 0, \quad i = 1, \ldots, m$$

- functions f_i now depend on a second variable u, the "uncertainty", which is constrained to lie in given set $\mathcal{U}$.
- Inherits convexity from nominal. Very tractable in some practically relevant cases.
- Complexity is high in general, but there are systematic ways to get relaxations.

Robust LP

Nominal problem:

$$\min_x \; c^\top x \; : \; a_i^\top x \le b_i, \;\; i = 1, \ldots, m.$$

Now assume that a_i is only known to belong to a given set $\mathcal{U}_i \subseteq \mathbb{R}^n$.

Robust counterpart:

$$\min_x \; c^\top x \; : \; \forall a_i \in \mathcal{U}_i, \;\; a_i^\top x \le b_i, \;\; i = 1, \ldots, m.$$

Robust LP

Uncertainty in the cost vector

Nominal problem:

$$\min_x \; c^\top x \;:\; a_i^\top x \le b_i, \quad i = 1, \ldots, m.$$

Now assume that the cost vector c is only known to belong to a given set $\mathcal{U} \subseteq \mathbb{R}^n$.

Robust counterpart:

$$\min_x \; \max_{c \in \mathcal{U}} \; c^\top x \;:\; a_i^\top x \le b_i, \quad i = 1, \ldots, m.$$

- We can extend the robust approach to cases with uncertainties affecting both cost vectors, coefficient matrix, and right-hand side vector b.
- Solution may be hard in general, but becomes easy for special uncertainty sets. We examine these special cases now.

A single inequality with uncertain coefficient vector

Next we examine the robust counterpart to a single inequality constraint:

$$\forall\, a \in \mathcal{U}, \;\; a^\top x \le b$$

where $\mathcal{U}$ takes the following forms:

- scenario uncertainty: $\mathcal{U}$ is a finite set of '''scenarios";
- $\mathcal{U}$ is a sphere, or more generally an ellipsoid;
- $\mathcal{U}$ is a box.

The above constraint can be written

$$b \ge \max_{a \in \mathcal{U}} \; a^T x.$$

Robust single inequality

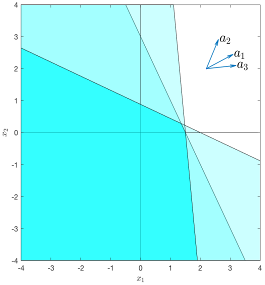
Scenario uncertainty

The scenario uncertainty model assumes that the coefficient vector a is only known to lie in a finite set in $\mathbb{R}^n$:

$$\mathcal{U} = \left\{ a^{(1)}, \ldots, a^{(K)} \right\},$$

with $a^{(k)} \in \mathbb{R}^n$ a "scenario", $k = 1, \ldots, K$. We have

$$\max_{a \in \mathcal{U}} a^\top x = \max_{1 \le k \le K} (a^{(k)})^\top x.$$



When $\mathcal{U}$ is a finite set of three scenarios, the set

$$\left\{ x \; : \; a^\top x \le b \; : \; \forall \, a \in \mathcal{U} \right\}$$

is a polyhedron made up of three half-spaces.

Robust single inequality

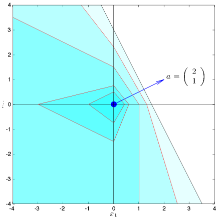
Box uncertainty

The box uncertainty model assumes that the coefficient vector a_i is only known to lie in a "box" (a hyper-rectangle in $\mathbb{R}^n$). In its simplest case, this uncertainty model has the form:

$$\mathcal{U} = \{a : \|a - \hat{a}\|_\infty \le \rho\} = \{a + \rho u : \|u\|_\infty \le 1\},$$

where $\rho \ge 0$ is a measure of the size of the uncertainty, and $\hat{a}$ is the nominal value of the coefficient vector. We have

$$\max_{a \in \mathcal{U}} a^\top x = \hat{a}^\top x + \rho \cdot \left(\max_{u : \|u\|_\infty \le 1} u^\top x \right) = \hat{a}^\top x + \rho \|x\|_1.$$



When $\mathcal{U}$ is a box, the set

$$\left\{x \ : \ a^T x \le b \ : \ \forall\, a \in \mathcal{U}\right\}$$

is a polyhedron, with 2^n vertices.

Robust single inequality

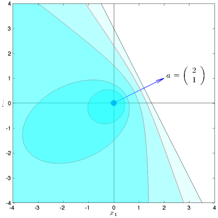
Spherical uncertainty

The spherical uncertainty model assumes that the coefficient vector a_i is only known to lie in a sphere. This uncertainty model has the form:

$$\mathcal{U} = \{a : \|a - \hat{a}\|_2 \le \rho\} = \{a + \rho u : \|u\|_2 \le 1\}.,$$

where $\rho \ge 0$ is a measure of the size of the uncertainty, and $\hat{a}$ is the nominal value of the coefficient vector. We have

$$\max_{a \in \mathcal{U}} a^\top x = \hat{a}^\top x + \rho \cdot \left(\max_{u : \|u\|_2 \le 1} u^\top x \right) = \hat{a}^\top x + \rho \|x\|_2.$$



When $\mathcal{U}$ is a sphere, the set

$$\left\{x : a^T x \le b : \forall\, a \in \mathcal{U}\right\}$$

is defined by a single SOCP constraint.

Robust single inequality

Ellipsoidal uncertainty

The ellipsoidal uncertainty model assumes that the coefficient vector a is only known to lie in a ellipse in $\mathbb{R}^n$.

This uncertainty model has the following form:

$$\mathcal{U} = \left\{a \ : \ (a - \hat{a})^\top P^{-1}(a - \hat{a}) \leq 1\right\},$$

where $\hat{a}$ represents the nominal value of the coefficient vector, and matrix $P = P^\top \succ 0$ determines the shape and size of the ellipse. Since $P \succ 0$, we can write $P = R^\top R$ for some matrix R. Then

$$\mathcal{U} = \{a = \hat{a} + Ru \ : \ \|u\|_2 \leq 1\},$$

and

$$\max_{a \in \mathcal{U}} a^\top x = \hat{a}^\top x + \max_{u \,:\, \|u\|_2 \leq 1} (Ru)^\top x = \hat{a}^\top x + \|R^\top x\|_2.$$

Robust LP with box uncertainty

The robust LP with box uncertainty:

$$\begin{array}{ll} \min\limits_{x} & c^\top x \\ \text{s.t.:} & \forall\, a_i \in \mathcal{B}_i \ : \ a_i^\top x \leq b_i \quad i = 1, \ldots, m, \end{array}$$

where $\mathcal{B}_i = \{\hat{a}_i + \rho_i u \ : \ \|u\|_\infty \leq 1\}$, $i = 1, \ldots, m$, is

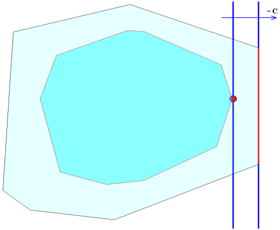
$$\begin{array}{ll} \min\limits_{x} & c^\top x \\ \text{s.t.:} & \hat{a}_i^\top x + \rho_i \|x\|_1 \leq b_i \quad i = 1, \ldots, m. \end{array}$$

This problem can in turn be expressed in standard LP form as

$$\begin{array}{lll} \min\limits_{x,u} & c^\top x & \\ \text{s.t.:} & \hat{a}_i^\top x + \rho_i \sum_{j=1}^n u_j \leq b_i, & i = 1, \ldots, m, \\ & -u_j \leq x_j \leq u_i, & j = 1, \ldots, n. \end{array}$$

Robust LP with box uncertainty

Geometry



In this example with box uncertainty, the robust counterpart's feasible set is inside the nominal feasible set, and has polyhedral boundaries; the robust problem is still an LP.

Robust LP with ellipsoidal uncertainty

The robust LP with ellipsoidal uncertainty:

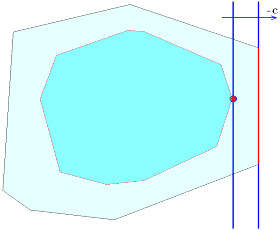
$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.:} \quad & \forall\, a_i \in \mathcal{E}_i \;:\; a_i^\top x \le b_i \quad i = 1, \ldots, m, \end{aligned}$$

where $\mathcal{E}_i = \{\hat{a}_i + R_i u \;:\; \|u\|_2 \le 1\}$, $i = 1, \ldots, m$, is the SOCP

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.:} \quad & \hat{a}_i^\top x + \|R_i^\top x\|_2 \le b_i \quad i = 1, \ldots, m. \end{aligned}$$

Robust LP with spherical uncertainty

Geometry



With spherical uncertainty, the robust counterpart's feasible set is inside the nominal feasible set, and has smooth boundaries, making the solution unique.

The nominal LP has many optimal points (red line), which means a solution might be very sensitive to data changes (such as if we change the direction of the objective slightly). In constrast, the solution to the robust LP is unique (red dot), irrespective of the choice of the objective. As a result, it not very sensitive to changes in the objective or other problem data.

Example

Drug production problem

Recall the drug production problem from lecture 10. The balance equation reads

$$0.01x_{\mathrm{RawI}} + 0.02x_{\mathrm{RawII}} - 0.05x_{\mathrm{DrugI}} - 0.600x_{\mathrm{DrugII}} = a^\top x \geq 0,$$

where $a_1 = 0.01$, $a_2 = 0.02$ contain the content of agent A (per kg) in each raw material, and a_3, a_4 contain the content of agent A in each of the drugs.

Uncertainty model: amount of active agent in raw material is uncertain:

$$a_1 \in [0.00995, 0.01005], \quad a_2 \in [0.0196, 0.0204],$$

representing a 0.5% and 2% box uncertainty around the nominal values.

Behavior of nominal and robust solutions

- If we disregard uncertainty in raw material's quality, and solve the nominal model, we obtain $x_{\mathrm{RawI}} = 0$, $x_{\mathrm{RawII}} = 438.79$, $x_{\mathrm{DrugI}} = 17,552$, $x_{\mathrm{DrugII}} = 0$, and profit $p^* = \$8819.66$.
- If the parameter for RawII takes the worst-case value, the nominal solution is not feasible. Decreasing x_{RawII} to make the constraint feasible again, leads to a 21% reduction in profit.
- Solving the robust counterpart instead, we get $x_{\mathrm{RawI}} = 877.73$, $x_{\mathrm{RawII}} = 0$, $x_{\mathrm{DrugI}} = 17,467$, $x_{\mathrm{DrugII}} = 0$, and profit $p^* = \$8294.56$. Profit is reduced by 5.95% only.

Chance-constrained LP

- Chance-constrained linear programs arise naturally from standard LPs, when some of the data describing the linear inequalities is uncertain and random.
- Consider an LP in standard inequality form:

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.:} \quad & a_i^\top x \le b_i, \quad i = 1, \ldots, m. \end{aligned}$$

- Suppose that the problem data vectors a_i, $i = 1, \ldots, m$, are not known precisely. Rather, all is known is that a_i are random vectors, with normal (Gaussian) distribution with mean value $\mathbb{E}\{a_i\} = \bar{a}_i$ and covariance matrix $\mathrm{var}\{a_i\} = \Sigma_i \succ 0$.
- In such a case, also the scalar $a_i^\top x$ is a random variable; precisely, it is a normal random variable with

$$\mathbb{E}\{a_i^\top x\} = \bar{a}_i^\top x, \quad \mathrm{var}\{a_i^\top x\} = x^\top \Sigma_i x.$$

Chance-constrained LP

- It makes no sense to impose a constraint of the form $a_i^\top x \leq b_i$, since the left-hand side of this expression is a normal random variable, which can assume any value, so such a constraint would always be violated by some outcomes of the random data a_i.
- We ask that the constraint $a_i^\top x \leq b_i$ be satisfied up to a given level of probability $p_i \in (0, 1)$.
- This level is chosen *a priori* by the user, and represents the probabilistic *reliability* level at which the constraint will remain satisfied in spite of random fluctuations in the data.
- The probability-constrained (or chance-constrained) counterpart of the nominal LP is therefore

$$\min_x \quad c^\top x \tag{1}$$

$$\text{s.t.:} \quad \mathrm{Prob}\{a_i^\top x \leq b_i\} \geq p_i, \quad i = 1, \ldots, m, \tag{2}$$

where p_i are the assigned reliability levels.

Chance-constrained LP

Proposition 1

Consider problem (1)–(2), under the assumptions that $p_i > 0.5$, $i = 1, \ldots, m$, *and that* a_i, $i = 1, \ldots, m$, *are independent normal random vectors with expected values* $\bar{a}_i$ *and covariance matrices* $\Sigma_i \succ 0$. *Then, (1)–(2) is equivalent to the SOCP*

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.:} \quad & \bar{a}_i^\top x \le b_i - \Phi^{-1}(p_i) \|\Sigma_i^{1/2} x\|_2, \quad i = 1, \ldots, m, \end{aligned} \tag{3}$$

where $\Phi^{-1}(p)$ *is the inverse cumulative probability distribution of a standard normal variable.*

Chance-constrained LP

Proof.

- We start by observing that

$$a_i^\top x \leq b_i \quad \Leftrightarrow \quad \frac{a_i^\top x - \bar{a}_i^\top x}{\sqrt{x^\top \Sigma_i x}} \leq \frac{b_i - \bar{a}_i^\top x}{\sqrt{x^\top \Sigma_i x}},$$

  where

$$\sigma_i(x) = \sqrt{x^\top \Sigma_i x} = \|\Sigma_i^{1/2} x\|_2.$$

- Defining

$$z_i(x) \doteq \frac{a_i^\top x - \bar{a}_i^\top x}{\sigma_i(x)}, \tag{4}$$

$$\tau_i(x) \doteq \frac{b_i - \bar{a}_i^\top x}{\sigma_i(x)}, \tag{5}$$

  we have that

$$\mathrm{Prob}\{a_i^\top x \leq b_i\} = \mathrm{Prob}\{z_i(x) \leq \tau_i(x)\}.$$

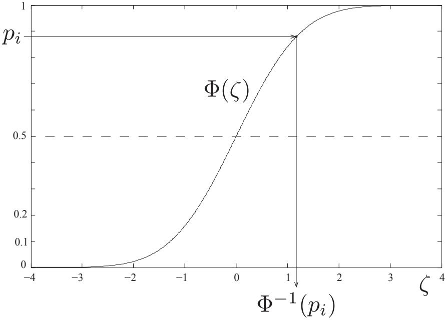
Chance-constrained LP

Proof (cnt).

- $z_i(x)$ is a standardized normal random variable (that is, a normal variable with zero mean and unit variance). Let $\Phi(\zeta)$ denote the standard normal cumulative probability distribution function, i.e.,

$$\Phi(\zeta) \doteq \mathrm{Prob}\{z_i(x) \leq \zeta\}.$$

Function $\Phi(\zeta)$ is well known and tabulated (also, it is related to the so-called *error function*, $\mathrm{erf}(\zeta)$, for which it holds that $\Phi(\zeta) = 0.5(1 + \mathrm{erf}(\zeta/\sqrt{2}))$).



Robust Least Squares

- Let us start from a standard LS problem:

$$\min_x \|Ax - y\|_2, \quad A \in \mathbb{R}^{m,n},\ y \in \mathbb{R}^m.$$

- Now assume that A is only known to be within a certain "distance" (in matrix space) to a given "nominal" matrix $\hat{A}$. Precisely, let us assume that

$$\|A - \hat{A}\| \leq \rho,$$

  where $\|\cdot\|$ denotes the largest singular value norm, and $\rho \geq 0$ measures the size of the uncertainty.
- Equivalently, we may say that $A = \hat{A} + \Delta$, where Δ is the uncertainty, which satisfies $\|\Delta\| \leq \rho$.
- We now address the *robust least-squares* problem:

$$\min_x \max_{\|\Delta\| \leq \rho} \|(\hat{A} + \Delta)x - y\|_2.$$

  The interpretation of this problem is that we aim at minimizing (with respect to x) the worst-case value (with respect to the uncertainty Δ) of the residual norm.

Robust Least Squares

- For fixed x, and using the fact that the Euclidean norm is convex, we have that

$$\|(\hat{A} + \Delta)x - y\|_2 \leq \|\hat{A}x - y\|_2 + \|\Delta x\|_2.$$

- By definition of the largest singular value norm, and given our bound on the size of the uncertainty, we have

$$\|\Delta x\|_2 \leq \|\Delta\| \cdot \|x\|_2 \leq \rho \|x\|_2.$$

  Thus, we have a bound on the objective value of the robust LS problem:

$$\max_{\|\Delta\| \leq \rho} \|(\hat{A} + \Delta)x - y\|_2 \leq \|\hat{A}x - y\|_2 + \rho \|x\|_2.$$

- The upper bound is actually attained by for

$$\Delta = \frac{\rho}{\|\hat{A}x - y\|_2 \cdot \|x\|_2} (\hat{A}x - y)x^\top.$$

Robust Least Squares

- Hence, the robust LS problem is equivalent to

$$\min_{x} \ \|\hat{A}x - y\|_2 + \rho\|x\|_2.$$

- This is a regularized LS problem, which can be cast in SOCP format as follows:

$$\begin{aligned} \min_{x,u,v} \quad & u + \rho v, \\ \text{s.t.} \quad & u \geq \|\hat{A}x - y\|_2, \\ & v \geq \|x\|_2. \end{aligned}$$