Short Course
Robust Optimization and Machine Learning

Lecture 5: Robust Optimization

Laurent El Ghaoui

EECS and IEOR Departments
UC Berkeley

Spring seminar TRANSP-OR, Zinal, Jan. 16-19, 2012

Outline

Outline

Optimization models

"Nominal" optimization problem:

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

f_0, f_i's are convex.

- Includes many problems arising in decision making, statistics.
- Efficient (polynomial-time) algorithms.
- Convex relaxations for non-convex problems.

Uncertainties are a pain!!

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.

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 chance counterpart

(Assume for simplicity there are no constraints)

$$\min_{x} \max_{p \in \mathcal{P}} \mathbf{E}_p f_0(x, u).$$

- Uncertainty is now random, obeys distribution p.
- Distribution p is only known to belong to a class $\mathcal{P}$ (*e.g.*, unimodal, given first and second moments).
- Complexity is high in general, but there are systematic ways to get relaxations.
- Rich variety of related models, including Value-at-Risk constraints.

In this lecture: our main goal is to introduce some important concepts in robust optimization, *e.g.* robust counterparts, affine recourse, distributional robustness.

Outline

Uncertainty models

Nominal problem:

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

We assume that $a_i = \hat{a}_i + \rho u_i$, where

- $\hat{a}_i$'s are the nominal coefficients.
- u_i's are the uncertain vectors, with $u_i \in \mathcal{U}_i$ but otherwise unknown.
- $\rho \ge 0$ is a measure of uncertainty.

Assumption that uncertainties affect each constraint independently is done without loss of generality.

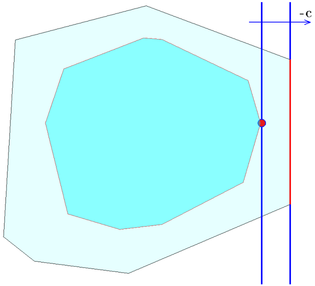
Robust counterpart

Robust counterpart:

$$\min_x \; c^T x \;:\; \forall u_i \in \mathcal{U}_i, \;\; (\hat{a}_i + \rho u_i)^T x \le b_i, \;\; i = 1, \ldots, m.$$

Solution may be hard, but becomes easy when:

- $\mathcal{U}_i$ are polytopic, given by their vertices ("scenarios");
- $\mathcal{U}_i$'s are "simple" sets such as ellipsoids, boxes, LMI sets, etc.
- Complexity governed by the support functions of sets $\mathcal{U}_i$.



Robust LP with ellipsoidal uncertainty.

Outline

Basic idea

Nominal LP:

$$\min_x \; c^T x \;:\; Ax \leq b.$$

We assume that A, b are affected by uncertainty in affine fashion. We assume uncertainty is available to “known by” some decision variables (*e.g.*, price revealed as time unfolds).

We seek an affinely adjusted robust solution (*i.e.*, a linear feedback).

Example

Nominal LP:

$$\min_x \; c^T x \;:\; Ax \leq b.$$

Assume that

- Right-hand side b is subject to uncertainty, $b(u) = \hat{b} + Bu$ with $u \in \mathcal{U}$.
- Decision variable can depend on (parts of) u: $x(u) = \hat{x} + Xu$.

Model information on u available to $x(\cdot)$ as $X \in \mathcal{X}$.

Affinely Adjustable Robust counterpart (AARC):

$$\min_{\hat{x}, X \in \mathcal{X}} \; \max_{u \in \mathcal{U}} \; c^T x(u) \;:\; \forall\, u \in \mathcal{U}, \;\; Ax(u) \leq b(u).$$

Above is tractable (provided $\mathcal{U}$ is).

Example

Assume $\mathcal{U} = [-\rho, \rho]^m$, we obtain the AARC

$$\min_{\hat{x}, X \in \mathcal{X}} c^T \hat{x} - \rho \|c^T X\|_1 \;:\; A\hat{x} + \rho s \leq \hat{b}, \quad s_i \geq \|e_i^T (AX - B)\|_1, \quad i = 1, \ldots, m.$$

We recover the “pure” robust counterpart with $X = 0$.

Case with coefficient uncertainty

Approach can be extended to cases when A, c are also uncertain.

- AARC is usually not tractable.
- Efficient approximations via SDP.

Outline

Chance constraints

Simple case

Consider an LP, and assume one of the constraints is $a^T x \leq b$, where $x \in \mathbf{R}^n$ is the decision variable.

If a is random, we can often deal with the chance constraint

$$\text{Prob}\left\{a^T x \leq b\right\} \geq 1 - \epsilon$$

easily. For example, if a is Gaussian with mean $\hat{a}$ and covariance matrix Γ, above is equivalent to

$$\hat{a}^T x + \kappa(\epsilon)\|\Gamma^{1/2} x\|_2 \leq b,$$

where $\kappa(\cdot)$ is a known function that is positive when $\epsilon < 0.5$.

More complicated chance constraints

Often, the random variable enters quadratically in the constraint. This happens for example when x includes affine recourse, and a depends linearly on some random variables.

We are led to consider

$$\text{Prob}\left\{(u,1)^T W(u,1) > 0\right\} \leq \epsilon$$

where W depends *affinely* on the decision variables. Above is hard, even in the Gaussian case.

Distributional robustness

Consider instead

$$\sup_{p \in \mathcal{P}} \mathbf{Prob}_p \left\{ (u,1)^T W (u,1) > 0 \right\} \leq \epsilon$$

where the sup is taken with respect to all distributions p in a specific class $\mathcal{P}$, specifying *e.g.*:

- Moments.
- Symmetry, unimodality.

Fact: when $\mathcal{P}$ is the set of distributions having zero mean and unit covariance, the condition $\sup_{p \in \mathcal{P}} P_{\rm wc} \leq \epsilon$ is equivalent to the LMI in M, v:

$$\mathbf{Tr}\, M \leq \epsilon v, \quad M \succeq 0, \quad M \succeq vJ + W,$$

where J is all zero but a 1 in the bottom-right entry.

Example

Transaction costs In many financial decision problems, the transaction costs can be modeled with

$$T(x, u) = \|A(x)u + b(x)\|_1,$$

for appropriate affine $A(\cdot), b(\cdot)$.

Example:

$$\sum_{t=1}^{T} |x_{t+1} - x_t|$$

with decision variable x_t an affine function of u.

This leads to consider quantities such as

$$\max_{u \sim (0, I)} \mathbf{E}\, T(x, u)$$

where $u \sim (0, I)$ refers to distributions with zero mean and unit covariance matrices.

A useful result

For given $m \times d$ matrix A and d-vector b, define

$$\phi := \max_{u \sim (0,I)} \mathbf{E} \, \|Au + b\|_1$$

Let a_i denote the i-th row of A ($1 \leq i \leq m$). Then

$$\frac{2}{\pi}\psi \leq \phi \leq \psi,$$

where

$$\psi := \sum_{i=1}^{m} \left\| \begin{pmatrix} a_i \\ b_i \end{pmatrix} \right\|_2 .$$

Note: ψ is convex in A, b, which allows to minimize it if A, b are affine in the decision variables.

Outline

Dynamic programming

- Finite-state, discrete-time Markov decision process.
- Finite-horizon control problem: minimize expected cost.
- $a \in \mathcal{A}$ denote actions, $s \in \mathcal{S}$ states, and $c_t(s, a)$ the cost for action a in state s at time t.

Bellman recursion (value iteration):

$$v_t(s) = \min_{a \in \mathcal{A}} \; c_t(s, a) + p_t(a)^T v_{t+1}, \quad s \in \mathcal{S}$$

with $p_t(a)$ the transition probabilities at time t under action a.

Uncertainty on transition matrix

We assume that *at each stage*, "nature" picks a transition probability vector $p_t(a)$ in a given set $\mathcal{P}_t(a)$.

Robust counterpart: the robust control problem, with "nature" the adversary.

Robust Bellman recursion:

$$v_t(s) = \min_{a \in \mathcal{A}} \; c_t(s,a) + \max_{p \in \mathcal{P}_t(a)} \; p^T v_{t+1}, \quad s \in \mathcal{S}.$$

For a wide variety of sets $\mathcal{P}_t(a)$, inner problem very easy to solve.

Entropy uncertainty model

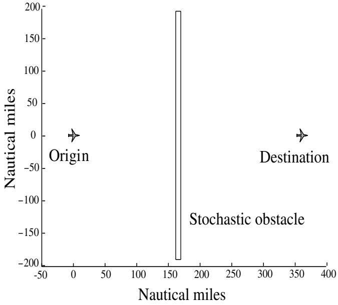
A natural way to model uncertainty in the transition matrices involves relative entropy bounds

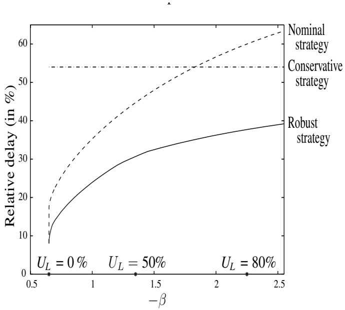
$$\mathcal{P} = \left\{ p \geq 0 \; : \; \sum_j p_j \log \frac{p_j}{q_j} \leq \beta, \; \sum_j p_j = 1 \right\}.$$

where $\beta > 0$ is a measure of uncertainty, and q is the nominal distribution.

The corresponding inner problem can be solved in $O(n)$ via bisection.

Example
Robust path planning





Outline

References

A. Bental, L. El Ghaoui, and A. Nemirovski.
Robust Optimization.
Princeton Series in Applied Mathematics. Princeton University Press, October 2009.

D. Bertsimas, D. Brown, and C. Caramanis.
Theory and applications of robust optimization.
SIAM Review, 2011.
To appear.

V. Guigues.
Robust production management.
Optimization Online, February 2011.
www.optimization-online.org/D8_HTML/2011/02/2935.html.

Arnab Nilim and Laurent El Ghaoui.
Robust control of Markov decision processes with uncertain transition matrices.
Oper. Res., 53(5):780–798, September-October 2005.