Lecture 24

Robust semidefinite programming: part one

- Principle
- Polytopic uncertainty
- SDP Relaxations
- Example: a combinatorial optimization problem

Principle

Recall **robust SDP:**

$$\text{minimize } c^T x \text{ subject to } \mathbf{F}_0 + x_1\mathbf{F}_1 + \cdots + x_m\mathbf{F}_m \succeq 0 \quad \text{for every } (\mathbf{F}_0, \ldots, \mathbf{F}_m) \in \mathcal{U}$$

what we'll do:

- we'll assume the set $\mathcal{U}$ has a "simple", *e.g.* **polytopic** structure (in the appropriate matrix space)
- get SDP approximation to original problem
- this SDP yields a bound (and a robust solution) to the original problem

later: generalize the approach to more general perturbation structures

Robust SDP with polytopic uncertainty

Assume $\mathcal{U}$ is a polytope that is described by its vertices

$$\mathcal{U} = \textbf{Co}\left\{(F_0^k, \ldots, F_l^k) \mid k = 1, \ldots, N\right\}$$

where $(F_0^k, \ldots, F_l^k)$ are l-tuples of symmetric matrices

Robust SDP is **equivalent** to **standard** SDP

$$\min\ c^T x \text{ subject to}$$
$$F^k(x) = F_0^k + \sum_{i=1}^{m} x_i F_i^k \succeq 0, \quad k = 1, \ldots, N$$

In this case the robust SDP is an SDP

hence, "tractable" ...

Really?

Complexity of polytopic model

complexity of previous SDP is linear in # vertices

but ...

- usually polytopes are described via linear inequalities
$$a_i^T x \le b_i, \quad i = 1, \ldots, l$$
- complexity of previous SDP is exponential in # linear inequalities

what we'll do: give up on accuracy to decrease complexity

Example: combinatorial optimization

a (NP-hard) quadratic problem

$$\phi := \max\ \delta^T W \delta \text{ subject to } \delta_i^2 = 1, \quad i = 1, \ldots, l,$$

where $W = W^T$ is given

WLOG, $W \succ 0$, so problem is equivalent to

$$\phi := \max\ \delta^T W \delta \text{ subject to } \|\delta\|_\infty \leq 1$$

(remember: $\|\delta\|_\infty := \max_{1\leq i \leq l} |\delta_i|$)

This is a robust SDP problem:

$$\phi = \min\ x \text{ subject to } x \geq \delta^T W \delta \text{ for every } \delta, \ \ \|\delta\|_\infty \leq 1$$

Polytopic approach

Describe uncertainty polytope by its vertices:

$$\{\delta \ | \ \|\delta\|_\infty \leq 1\} = \textbf{Co}\ \{\delta_1, \ldots, \delta_N\}$$

where $N = 2^l$

We have to solve

$$\phi = \min\ x \text{ subject to } x \geq \delta_i^T W \delta_i, \ \ i = 1, \ldots, N = 2^l$$

This is an SDP ...with **exponential** # of constraints!

Later: we will see an **approximation** procedure to this problem that leads to an SDP

(with complexity polynomial in l)

Robust SDP with affine uncertainty

nominal SDP:

$$\min\ c^T x \text{ subject to } F(x) = F_0 + \sum_{i=1}^{m} x_i F_i \succeq 0$$

where F_i's are symmetric

uncertainty model: assume $\mathbf{F}_i$ is uncertain

$$\mathbf{F}_i(\delta) = F_i^0 + \sum_{j=1}^{l} \delta_j F_i^j$$

where parameter vector $\delta \in \mathbf{R}^l$ is unknown, but componentwise bounded:

$$|\delta_j| \leq 1, \quad j = 1, \ldots, l$$

Let

$$\mathbf{F}(x,\delta) = \mathbf{F}_0(\delta) + \sum_{i=1}^{m} x_i \mathbf{F}_i(\delta)$$

robust SDP:

$$\min c^T x \quad : \ \mathbf{F}(x,\delta) \succeq 0 \text{ for every } \delta \in \mathbf{R}^l, \ \ \|\delta\|_\infty \leq 1$$

Matrix representation

fact: we can write

$$\mathbf{F}(x,\delta) = F(x) + L(x)\Delta R + (L(x)\Delta R)^T$$

where

- symmetric $F(x)$ and $L(x)$ are affine in x
- R fixed matrix
- Δ is diagonal matrix with repeated elements

  $$\Delta = \mathbf{diag}\ (\delta_1 I_{r_1}, \ldots \delta_l I_{r_l})$$

  with $r_1, \ldots, r_l$ appropriate integers

Bound on δ translates as

$$\|\Delta\| \leq 1$$

Define uncertainty set

$$\boldsymbol{\Delta} = \{\Delta = \mathbf{diag}\,(\delta_1, \ldots, \delta_l) \ \mid \|\Delta\| \leq 1\}$$

Simplifying assumption: we'll assume $r_i = 1$, *i.e.* Δ is an arbitrary diagonal matrix (bounded by 1 in norm)

Robustness analysis

Fix x, seek to enforce the robustness condition

$$\mathbf{F}(x,\Delta) \succeq 0 \text{ for every } \Delta \in \boldsymbol{\Delta}$$

i.e.,

$$\begin{aligned}\mathbf{F}(x,\Delta) &= F(x) - L(x)L(x)^T - R^T R \\ &\quad + \left[\, L \;\; R^T \,\right] \left[\begin{array}{cc} I & \Delta \\ \Delta^T & I \end{array}\right] \left[\begin{array}{c} L^T \\ R \end{array}\right] \succeq 0,\end{aligned}$$

for every diagonal Δ, with norm bound $\|\Delta\| \leq 1$, *i.e.* such that

$$\left[\begin{array}{cc} I & \Delta \\ \Delta^T & I \end{array}\right] \succeq 0$$

(obvious) **sufficient condition:**

$$F(x) \succeq L(x)L(x)^T + R^T R$$

Probably very conservative: we have not used fact that Δ is constrained to be diagonal

Using invariance

previous condition can be refined by noting that the LFR of $\mathbf{F}$ is **invariant** via (diagonal) scaling

if $S \succ 0$ is a given diagonal matrix, then the scaling

$$L(x) \to L(x)S^{-1/2}, \quad R \to S^{1/2}R$$

leaves the LFR of $\mathbf{F}(x,\cdot)$ invariant

we obtain the sufficient condition for robustness:

$$\exists\ S \succ 0, S \text{ diagonal}, \; F(x) \succeq L(x)S^{-1}L(x)^T + R^T S R$$

(equivalent to an LMI in x, S ...)

LMI form

Under some technical condition, previous (sufficient) robustness condition

$$\exists\ S \succ 0, S \text{ diagonal, } F(x) \succeq L(x)S^{-1}L(x)^T + R^TSR$$

is equivalent to LMI in **both** x and S

$$\begin{bmatrix} F(x) - R^TSR & L(x) \\ L(x)^T & S \end{bmatrix} \succeq 0$$

interpretation:

- The above LMI represents a convex set in the space of (x, S) variables
- The projection of this set on the space of x-variables is an inner approximation of the set of robustly feasible solutions

$$\{x \mid \mathbf{F}(x, \Delta) \succeq 0 \text{ for every } \Delta \in \mathbf{\Delta}\}$$

note: try to derive technical condition!

SDP approximation to robust SDP

We obtain that an **upper bound** on the robust SDP

$$\min_x c^Tx \quad : \ \mathbf{F}(x, \delta) \succeq 0 \text{ for every } \delta \in \mathbf{R}^l, \ \ \|\delta\|_\infty \leq 1$$

is given by the standard SDP

$$\min_{S,x} c^Tx \quad : \ S \text{ diagonal,} \quad \begin{bmatrix} F(x) - R^TSR & L(x) \\ L(x)^T & S \end{bmatrix} \succeq 0$$

SDP relaxation to combinatorial optimization problem

remember the (NP-hard) quadratic problem

$$\phi := \max\ \delta^T W \delta \text{ subject to } \delta_i^2 \le 1, \quad i = 1, \ldots, l,$$

where $W = W^T$ is given (WLOG, $W \succ 0$)

This is a robust SDP problem with affine uncertainty and maximum-norm bounds:

$$\phi = \min x \ \ \mathbf{F}(x, \Delta) \succeq 0 \text{ for every } \Delta \in \mathbf{\Delta},$$

where

$$\mathbf{F}(x, \Delta) = \begin{bmatrix} x & \delta^T \\ \delta & W^{-1} \end{bmatrix},$$

and

$$\mathbf{\Delta} = \{\Delta = \mathbf{diag}(\delta_1, \ldots, \delta_l) \mid \|\Delta\| \le 1\}$$

We've seen that polytopic uncertainty is useless here (does not reduce exponential complexity)

A close look at a combinatorial problem

Let us apply the relaxation outlined before

We have:

$$\mathbf{F}(x, \Delta) = \begin{bmatrix} x & \delta^T \\ \delta & W^{-1} \end{bmatrix} = F(x) + L\Delta R + (L\Delta R)^T,$$

with $\Delta = \textbf{diag}\,(\delta_1, \ldots, \delta_l)$, and

$$F(x) = \begin{bmatrix} x & 0 \\ 0 & W^{-1} \end{bmatrix}, \quad R = \left[\begin{array}{c|c} 1 & \\ \vdots & 0 \\ 1 & \end{array}\right], \quad L = \begin{bmatrix} 0 \\ I \end{bmatrix}$$

Previous approach yields robust SDP approximation

$$\begin{array}{ll} \min_{S,x} x & : S \text{ diagonal}, \\ & \begin{bmatrix} F(x) - R^T S R & L(x) \\ L(x)^T & S \end{bmatrix} \succeq 0 \end{array}$$

Since $R^T S R = \textbf{diag}\,(\textbf{Tr}\,S, 0, \ldots, 0)$, rewrite above as

$$\begin{array}{ll} \min_{S,x} x & : S \text{ diagonal}, \\ & \left[\begin{array}{c|c} \begin{bmatrix} x - \textbf{Tr}\,S & 0 \\ 0 & W^{-1} \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ \hline \begin{bmatrix} 0 & I \end{bmatrix} & S \end{array}\right] \succeq 0 \end{array}$$

Upper bound result

Previous inequality implies $S \succ 0$ and can be rewritten as

$$x \geq \mathbf{Tr}\, S \text{ and } W^{-1} \succeq S^{-1}$$

Minimizing x is thus equivalent to

$$\min \mathbf{Tr}\, S \text{ subject to } S \text{ diagonal}, W \preceq S$$

features of above problem

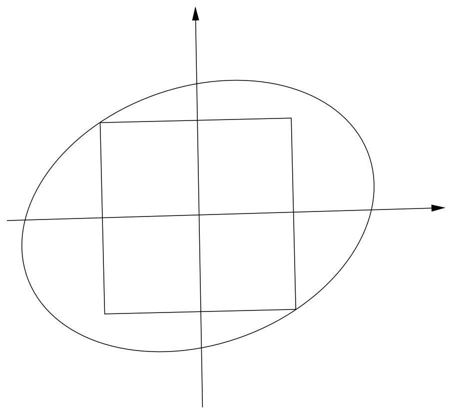
- an SDP in scalar variable and diagonal S
- polynomial complexity in l (size of original problem)
- provides best bounds known so far

Geometric interpretation

Define level sets associated to matrix W:
$\{\delta \mid \delta^T W \delta \leq x\}$

(Since $W \succ 0$, these are ellipsoids)

We seek the smallest level x such that corresponding level set containts unit hypercube



For arbitrary W, this is a hard problem (must enumerate all vertices)

Using intermediary ellipses

However, problem is **easy** if level sets are parallel to axes

Ellipsoids parallel to the axes are parametrized by diagonal matrices:

$$\mathcal{E}_S = \{\delta \mid \delta^T S \delta \leq 1\}$$

where $S \succ 0$ is **diagonal**

If $W = S$, where $S \succ 0$ is diagonal, then **any** vertex is optimal; taking all components of δ equal to one leads to optimal objective

$$\phi = \mathbf{Tr}\, S$$

Relaxation procedure: optimize over S such that

- $\mathcal{E}_S$ contains the unit hypercube
  equivalent to
- $\mathcal{E}_S$ is contained by the level set
  equivalent to $W \preceq S$

Dual problem

The problem dual to

$$\min \mathbf{Tr}\, S \text{ subject to } S \text{ diagonal}, W \preceq S$$

is

$$\max\ \mathbf{Tr}\, WX \text{ subject to } X \succeq 0,\ \ X_{ii} = 1,\ \ 1 \leq i \leq n$$

Interpretation: with $X := \delta\delta^T$, original problem:

$$\phi := \max\ \delta^T W \delta \text{ subject to } \delta_i^2 \leq 1, \quad i = 1, \ldots, l,$$

can be written as

$$\max\ \mathbf{Tr}\, WX \text{ subject to } X \succeq 0,\ \ X_{ii} \leq 1,\ \ 1 \leq i \leq n, \quad \mathbf{Rank} X = 1$$

(we've used the fact that $\delta^T W \delta = \mathbf{Tr}\, W\delta\delta^T$.)

SDP relaxation amounts to relax the rank constraint

Primalization

SDP relaxation is the most efficient currently available (in terms of closeness to the actual optimum, in a certain stochastic sense, see Goemans and Williamson's, 1994, Nesterov, 1996)

idea of proof: find efficient *primalization* technique, which computes a suboptimal set of parameters δ from the solution X of the (dual) SDP

$$\max \ \mathbf{Tr}\, WX \text{ subject to } X \succeq 0, \quad X_{ii} = 1, \quad 1 \leq i \leq n$$

then show that the suboptimum is "close" to upper bound

original primalization technique is stochastic:

- primalization obtained by random choice
- compare expected value of suboptimum with SDP upper bound

Unstructured matrix perturbations

Sometimes, it is useful to consider case when uncertain model is

$$\mathbf{F}(x, \delta) = F(x) + L(x)\Delta R + (L(x)\Delta R)^T$$

where Δ is bounded in norm, but otherwise **arbitrary**

(we've looked at the case when Δ was restricted to be diagonal)

We can apply the same technique; the results turn out to be exact

Robustness analysis

Fix x, seek to enforce the robustness condition

$$\mathbf{F}(x,\Delta) \succeq 0 \text{ for every } \Delta, \quad \|\Delta\| \leq 1$$

i.e.,

$$\begin{aligned} \mathbf{F}(x,\Delta) &= F(x) - L(x)L(x)^T - R^T R \\ &\quad + \begin{bmatrix} L & R^T \end{bmatrix} \begin{bmatrix} I & \Delta \\ \Delta^T & I \end{bmatrix} \begin{bmatrix} L^T \\ R \end{bmatrix} \succeq 0, \end{aligned}$$

for every Δ, with norm bound $\|\Delta\| \leq 1$, *i.e.* such that

$$\begin{bmatrix} I & \Delta \\ \Delta^T & I \end{bmatrix} \succeq 0$$

(obvious) **sufficient condition:**

$$F(x) \succeq L(x)L(x)^T + R^T R$$

can we refine this? (Δ not constrained to be diagonal anymore)

Using invariance

Turns out that some invariance property still holds

previous condition can be refined by noting that the LFR of $\mathbf{F}$ is **invariant** via **scalar** scaling

if $\tau > 0$ is a given scalar, then the scaling

$$L(x) \to \frac{1}{\sqrt{\tau}} L(x), \quad R \to \sqrt{\tau} R$$

leaves the LFR of $\mathbf{F}(x,\cdot)$ invariant

we obtain the sufficient condition for robustness:

$$\exists\ \tau > 0, F(x) \succeq \frac{1}{\tau} L(x)L(x)^T + \tau R^T R$$

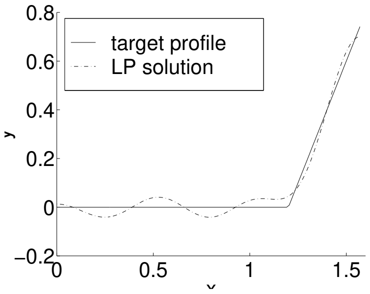
equivalent to an LMI in x, τ:

$$\begin{bmatrix} F(x) - \tau R^T R & L(x) \\ L(x)^T & \tau \cdot I \end{bmatrix} \succeq 0$$

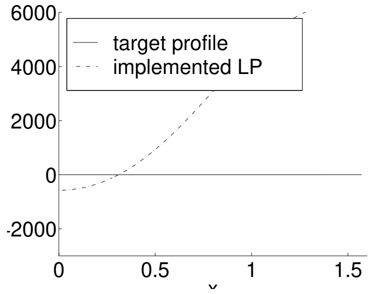
fact: above approximation is **exact**

Example: robust antenna array design

Recall from lecture 1:



linear programming solution
$\min \|Ax - b\|_\infty$

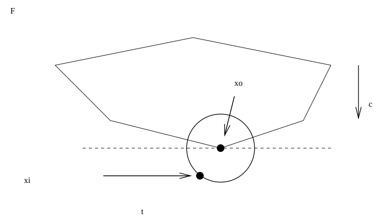


implemented solution
(relative error in $x \leq .1\%$)

what went wrong?

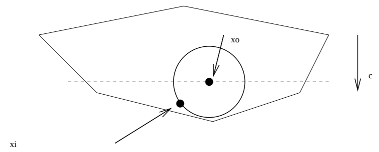
Antenna array design via robust SDP

what went wrong is that optimal solution computed from LP has **large** magnitude



hence, even small **relative** errors are large in the absolute sense!

idea: make sure design is implementable



robust implementability requires that solutions at given relative distance must be in feasible set

LP with robust implementability

Consider a linear (affine) inequality constraint

$$a^T x \le b$$

and seek x that is *robustly implementable*, that is such that

$$a^T(x + \delta x) \le b \text{ whenever } \|\delta x\|_2 \le \rho \|x\|_2$$

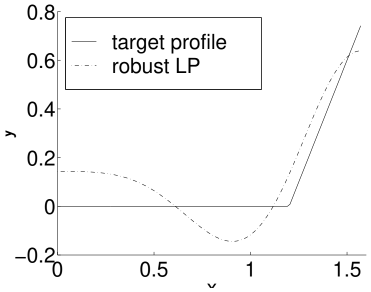
where $\rho \ge 0$ is given

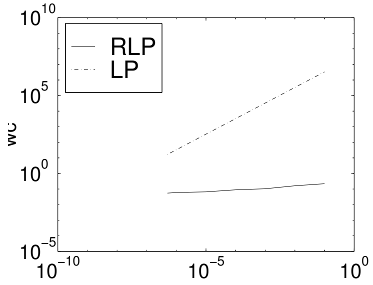
rewrite this as a robust SDP constraint with unstructured uncertainty:

$$a^T(I + \Delta)x \le b \text{ for every } \Delta, \quad \|\Delta\| \le \rho$$

Above can be written down exactly as a standard LMI (involving original design variable x and a new scalar variable τ)

Robust vs. non robust antenna array design





performance	nominal	perturbed (0.1%)
LP	0.04	10^4
RLP	0.14	0.15

morale: with little extra computer cost, obtain a design that is implementable **and** performant

What's next

We've looked at at special case of robust SDP:

- "unstructured", matrix uncertainty
- OR
  - data depends affinely on uncertain vector δ
  - δ is bounded componentwise

next, we'll generalize our results:

- nonlinear (algebraic) dependence of data on uncertainty
- other kinds of norm bounds (*e.g.*, Euclidean)