ORF 523	LECTURE 16	Spring 2017, Princeton University
Instructor: A.A. Ahmadi		
Scribe: G. Hall		Tuesday, May 2, 2017

When in doubt on the accuracy of these notes, please cross check with the instructor's notes, on `aaa.princeton.edu/orf523`*. Any typos should be emailed to gh4@princeton.edu.*

In this lecture, we give a brief introduction to

- robust optimization (Section 1)
- robust control (Section 2).

1 Robust optimization

"To be uncertain is to be uncomfortable, but to be certain is to be ridiculous."

Chinese proverb [1].

- So far in this class, we have assumed that an optimization problem is of the form

$$\begin{aligned} &\min_{x}.\ f(x) \\ &g_i(x) \leq 0,\ i = 1, \ldots, n, \\ &h_j(x) = 0,\ j = 1, \ldots, m, \end{aligned}$$

  where f, g_i, h_j are exactly known. In real life, this is most likely not the case; the objective and constraint functions are often not precisely known or at best known with some noise.

- *Robust optimization* is an important subfield of optimization that deals with uncertainty in the data of optimization problems. Under this framework, the objective and constraint functions are only assumed to belong to certain sets in function space (the so-called "uncertainty sets"). The goal is to make a decision that is feasible no matter what the constraints turn out to be, and optimal for the worst-case objective function.

1.1 Robust linear programming

- In this section, we will be looking at the basic case of robust linear programming. We will consider two types of uncertainty sets: polytopic and ellipsoidal.

- A robust LP is a problem of the form:

$$\begin{aligned} &\min_{x}. \ c^T x \\ &\text{s.t. } a_i^T x \le b_i, \ \forall a_i \in U_{a_i}, \ \forall b_i \in U_{b_i}, i = 1, \ldots, m, \end{aligned} \tag{1}$$

  where $U_{a_i} \subseteq \mathbb{R}^n$ and $U_{b_i} \subseteq \mathbb{R}$ are given uncertainty sets.

- Notice that with no loss of generality we are assuming that there is no uncertainty in the objective function. This is because of the following equations

$$\begin{aligned} &\min_{x}. \ \max_{c \in U_c}. \ c^T x \\ &\text{s.t. } a_i^T x \le b_i, \ \forall a_i \in U_{a_i}, \ \forall b_i \in U_{b_i}, \ i = 1, \ldots, m. \\ &\qquad \Updownarrow \\ &\min_{x,\alpha}. \ \alpha \\ &c^T x \le \alpha, \ \forall c \in U_c \\ &a_i^T x \le b_i, \ \forall a_i \in U_{a_i}, \ \forall b_i \in U_{b_i}, \ i = 1, \ldots, m. \end{aligned}$$

1.1.1 Robust LP with polytopic uncertainty

- This is the special case of the previous problem where U_{a_i} and U_{b_i} are polyhedra; i.e.,

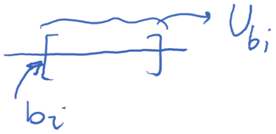
$$U_{a_i} = \{a_i | \ D_i a_i \le d_i\},$$

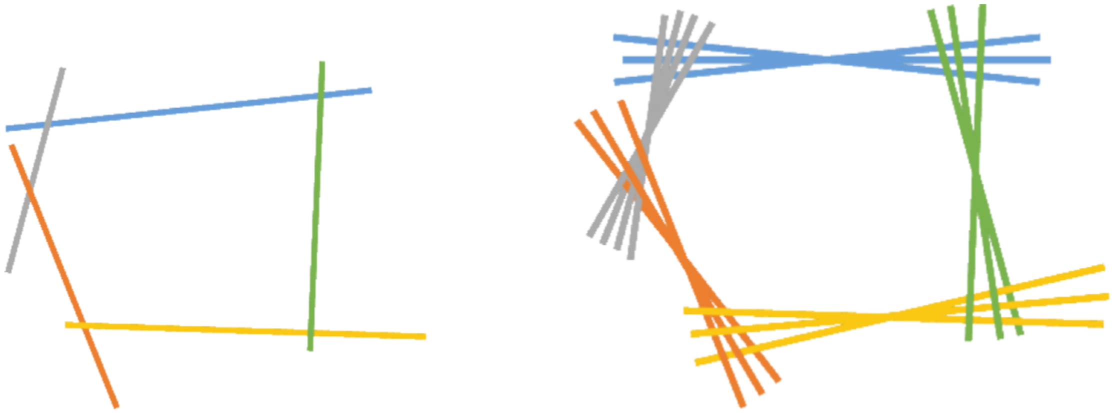
  where $D_i \in \mathbb{R}^{k_i \times n}$ and $d_i \in \mathbb{R}^{k_i}$ are given to us as input. Similarly, each U_{b_i} is a given interval in $\mathbb{R}$.

- Clearly, we can get rid of the uncertainty in b_i because the worst-case scenario is achieved at the lower end of the interval. So our problem becomes

$$\begin{aligned} &\min_{x}. \ c^T x \\ &\text{s.t. } a_i^T x \le b_i, \ \forall a_i \in U_{a_i}, \ i = 1, \ldots, m, \end{aligned} \tag{2}$$

where $U_{a_i} = \{a_i | \; D_i a_i \leq d_i\}$ With some abuse of notation, we are reusing b_i to denote the lower end of the interval:





(a) Feasible set of an LP with no uncertainty (b) Feasible set of an LP with polytopic uncertainty

The linear program (2) can be equivalently written as

$$\begin{aligned} &\min_x . \; c^T x \\ &\text{s.t.} \; \begin{bmatrix} \max\limits_{a_i} . \; a_i^T x \\ D_i a_i \leq d_i \end{bmatrix} \leq b_i, \; i = 1, \ldots, m. \end{aligned} \tag{3}$$

Our strategy will be to change the min-max problem to a min-min problem to combine the two minimization problems. To do this, we take the dual of the inner optimization problem in (3), which is given by

$$\begin{aligned} &\min_{p_i \in \mathbb{R}^{k_i}} . \; p_i^T d_i \\ &D_i^T p_i = x \\ &p_i \geq 0. \end{aligned}$$

By strong duality, both problems have the same optimal value so we can replace (3) by

$$\begin{aligned} &\min_{x}.\ c^T x \\ &\text{s.t.}\ \begin{bmatrix} \min\limits_{p_i \in \mathbb{R}^{k_i}}.\ p_i^T d_i \\ D_i^T p_i = x \\ p_i \geq 0 \end{bmatrix} \leq b_i,\ \ i = 1, \ldots, m. \end{aligned} \tag{4}$$

But this is equivalent to

$$\begin{aligned} \min_{x, p_i}.\ \ & c^T x \\ \text{s.t.}\ \ & p_i^T d_i \leq b_i,\ \ i = 1, \ldots, m \\ & D_i^T p_i = x,\ \ i = 1, \ldots, m, \\ & p_i \geq 0,\ \ i = 1, \ldots, m. \end{aligned} \tag{5}$$

This equivalence is very easy to see: suppose we have an optimal x, p for (5). Then x is also feasible for (4) and the objective values are the same. Conversely, suppose we have an optimal x for (4). As x is feasible for (4), there must exist p verifying the inner LP constraint. Hence, (x, p) would be feasible for (5) and would give the same optimal value.

Duality has enabled us to solve a robust LP with polytopic uncertainty just by solving a regular LP.

1.1.2 Robust LP with ellipsoidal uncertainty

We consider again an LP of the form (2) (i.e., no uncertainty in b_i), but this time we have

$$U_{a_i} = \{\bar{a}_i + P_i u |\ ||u||_2 \leq 1\}, i = 1, \ldots, m,$$

where $P_i \in \mathbb{R}^{n \times n}$ and $\bar{a}_i \in \mathbb{R}^n,\ i = 1, \ldots, m$ are part of the input.

- The sets U_{a_i} are ellipsoids, which gives the name *ellipsoidal uncertainty* to this type of uncertainty.
- If $P_i = I$, then the uncertainty sets are exactly spheres.
- If $P_i = 0$, then a_i is fixed, and there is no uncertainty.

Once again, we can formulate our problem as

$$\begin{aligned} &\min_{x}.\ c^T x \\ &\text{s.t. } \begin{bmatrix} \max_{a_i}.\ a_i^T x \\ a_i \in U_{a_i} \end{bmatrix} \leq b_i, i = 1, \dots, m. \end{aligned} \tag{6}$$

This time, the interior maximization problem has an explicit solution, which makes the problem easier. Indeed,

$$\begin{aligned} \max\{a_i^T x |\ a_i \in U_{a_i}\} &= \bar{a}_i{}^T x + \max\{u^T P_i^T x |\ ||u||_2 \leq 1\} \\ &= \bar{a}_i{}^T x + ||P_i^T x||_2, \end{aligned}$$

where the last equality is due to Cauchy-Schwarz applied to u and $P_i^T x$. Then, problem (6) can be rewritten as

$$\begin{aligned} &\min_{x}.\ c^T x \\ &\text{s.t. } \bar{a}_i{}^T x + ||P_i^T x||_2 \leq b_i \end{aligned}$$

which is an SOCP! Hence, a robust LP with ellipsoidal uncertainty can be solved efficiently by solving a single SOCP.

1.2 Robust SOCP with ellipsoidal uncertainty

Robust optimization is not restricted to linear programming. Many results are available for robust counterparts of other convex optimization problems with various types of uncertainty sets. For example, the robust counterpart of an uncertain SOCP (and hence an uncertain convex QCQP) with ellipsoidal uncertainty sets can be formulated as an SDP [3, Section 4.5].

Unfortunately, the robust counterpart of convex optimization problems does not always turn out to be a tractable problem. For example, the robust counterpart of an SOCP with polyhedral uncertainty is NP-hard [5], [2], [4]. Similarly, the robust counterpart of SDPs with pretty much any type of uncertainty is NP-hard. For example, even the following basic question is NP-hard [7]: given lower and upper bounds on entries of a matrix $l_{ij} \leq A_{ij} \leq u_{ij}$, is it true that all matrices in the family are positive semidefinite?
A good survey on tractability of robust counterparts of convex optimization problems is by Bertsimas et al. [5].

2 Robust stability of linear systems

In this section, we present one of the most basic and fundamental problems in robust control, namely, the problem of deciding robust stability of a linear system. Recall from our previous lectures that given a matrix $A \in \mathbb{R}^{n \times n}$, the linear dynamical system

$$x_{k+1} = Ax_k,$$

is globally asymptotically stable (GAS) if and only if

$$\rho(A) < 1,$$

where $\rho(A)$ is the spectral radius of A. We also saw that this is the case if and only if

$$\exists P \succ 0 \text{ s.t. } A^T PA \prec P.$$

For simplicity, let us call a matrix A with $\rho(A) < 1$ a *stable* matrix.
We now consider a related problem: we would like to study the stability of a linear system but when the matrix A is not exactly known. A common model for accounting for uncertainty in A is the following. We assume we know that

$$A \in \mathcal{A} := \text{conv}(A_1, \dots, A_m), \tag{7}$$

where $A_1, \dots, A_m$ are given $n \times n$ matrices. If all matrices $A \in \mathcal{A}$ are stable, then the system is said to be *robustly stable.*

Example: A population model of Sumatran tigers

Source: worldwildlife.org

A team of biologists has established that the growth dynamic of the population of Sumatran tigers is described by the following model:

$$\begin{pmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ \vdots \\ x_{k+1}^n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & 0 & 0 & \dots & 0 \\ 0 & a_{32} & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & a_{n,n-1} & 0 \end{pmatrix} \begin{pmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ \vdots \\ x_k^n \end{pmatrix}$$

In this model, the population is divided into n age groups in increasing order. The variable x_k^i denotes the number of individuals in age group i (e.g., 5-10 years) during the k^{th} year. The dynamic equation given above relates the number of individuals alive (in each age group) during year $(k+1)$ to the number of individuals alive in year k. The structure of A is intuitive: at each time stage, only a fraction $a_{(i+1)i}$ of people in age group i make it to the age group $i+1$. At the same time, each age group i contributes a fraction a_{1i} to the newborns in the next stage. Given these dynamics, the biologists would like to determine whether it is likely that this particular breed of tigers will go extinct.

This is a usual linear system of the type $x_{k+1} = Ax_k$. If the matrix A was perfectly known to the scientists, then they would be able to determine whether the tigers would go extinct, based on the spectral radius of A. However, in this case, the biologists have two different estimates of A at their disposal, denoted by A_1 and A_2, which come from two different teams of field biologists (one team is in Sumatra and one in Borneo). As both teams usually produce reliable work, they do not know which matrix to use for their computations and wonder if the following is true: if A_1 and A_2 are stable, is $\theta A_1 + (1-\theta)A_2$ stable $\forall \theta \in [0,1]$? In other words, is the system robustly stable?

The answer is actually negative! Stability of $A_1, \dots, A_m$ does not imply robust stability; i.e., stability of the convex hull in (7). This can be seen in the following example. Consider

$$A_1 = \begin{pmatrix} 0.2 & 0.3 & 0.7 \\ 0.9 & 0 & 0 \\ 0 & 0.8 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.3 & 0.9 & 0.4 \\ 0.5 & 0 & 0 \\ 0 & 0.9 & 0 \end{pmatrix},$$

we have $\rho(A_1) = 0.9887 < 1$ and $\rho(A_2) = 0.9621 < 1$, so both matrices are stable. But if we take $\theta = \frac{3}{5}$ then

$$\frac{3}{5}A_1 + \frac{2}{5}A_2 = \begin{pmatrix} 0.24 & 0.54 & 0.58 \\ 0.74 & 0 & 0 \\ 0 & 0.84 & 0 \end{pmatrix}$$

is not stable as it has spectral radius $\rho = 1.0001 > 1$.

In fact, determining when the system is robustly stable is NP-hard (see [7]). However, there are efficiently checkable sufficient conditions for this property.

Lemma 1. *Let $A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$. If there exists a matrix $P \succ 0$ s.t.*

$$A_i^T P A_i \prec P,\ \ i = 1, \ldots, m, \tag{8}$$

then $\rho(A) < 1, \forall A \in \mathcal{A}$.

Proof: Let

$$A = \sum_i \alpha_i A_i,$$

where $\alpha_i \geq 0, i = 1, \ldots, m$ and $\sum_i \alpha_i = 1$. If there exists $P \succ 0$ such that

$$A_i^T P A_i \prec P, \forall i = 1, \ldots m,$$

then, by taking the Schur complement, we get the LMIs

$$\left[\begin{array}{c|c} P & A_i^T \\ \hline A_i & P^{-1} \end{array}\right] \succ 0, i = 1, \ldots, m.$$

Multiplying by $\alpha_i \geq 0$ on both sides and summing, we get

$$\left[\begin{array}{c|c} P & A^T \\ \hline A & P^{-1} \end{array}\right] \succ 0,$$

which implies that $P \succ 0$ and $A^T P A \prec P$, using the Schur complement again. Hence, A is stable. □

Note that the LMIs in (8) are sufficient but not necessary for robust stability. There are indeed better LMI-based sufficient conditions for robust stability in the literature.

Notes

Further reading for this lecture can include the survey paper on robust optimization in [5] or an early paper on the topic by some people you know [6] ;)

References

[1] A. Ben-Tal and L. El-Ghaoui. *Robust Optimization.* Princeton University Press, 2009.

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

[3] A. Ben-Tal and A. Nemirovski. Robust optimization–methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

[4] A. Ben-Tal, A. Nemirovski, and C. Roos. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal on Optimization*, 13(2):535–560, 2002.

[5] D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

[6] J.M. Mulvey, R. J. Vanderbei, and S. A. Zenios. Robust optimization of large-scale systems. *Operations Research*, 43(2):264–281, 1995.

[7] A. Nemirovski. Several NP-hard problems arising in robust stability analysis. *Math. Control Signals Systems*, 6:99–105, 1993.