ECE285: Semidefinite and sum-of-squares optimization **Winter 2023**

Lecture 17: Robust optimization

Lecturer: Yang Zheng *Scribe: Yang Zheng*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. Any typos should be sent to* `zhengy@eng.ucsd.edu`*.*

Learning goals:

1. Robust optimization
2. Robust LP
3. Robust SOCP/(convex) QCQP
4. Robust SDP

1 Robust Optimization

In this course, we have focused on an optimization problem of the from

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0,\ i = 1, \ldots, m, \\ & h_j(x) = 0,\ j = 1, \ldots, p, \end{aligned} \tag{1}$$

where the functions f, g_i, h_j are exactly known. However, in practice, they always come from some real applications, and the exact functions are often not precisely known or at best known with some noise.

Robust optimization is an important sub-field of optimization, which deals with uncertainty in the data of optimization problems. In this framework, the objective and constraints are assumed to belong to certain sets. In particular, we consider the minimization of an objective $f(x)$ subject to constraints $g_i(x, u_i) \leq 0$ with uncertain parameters u_i. The general formulation of robust optimization is

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & g_i(x, u_i) \leq 0, \quad \forall u_i \in \mathcal{U}_i, \ \ i = 1, \ldots, m, \end{aligned} \tag{2}$$

where $x \in \mathbb{R}^n$ is a vector of decision variables, $f, g_i : \mathbb{R}^n \to \mathbb{R}$ are functions, and the uncertainty parameters $u_i \in \mathbb{R}^k$ are assumed to take arbitrary values in the uncertainty sets $\mathcal{U}_i \subset \mathbb{R}^k$. The goal of (2) is to find a solution with the minimum cost among all those solutions that are feasible for all realization of the uncertainty $u_i \in \mathcal{U}_i$. Some comments on (2) are

- If $\mathcal{U}_i$ only contain one single point, then the corresponding constraint has no uncertainty.
- If all the uncertainty sets $\mathcal{U}_i$ contain finite discrete points, then (2) is the same as (1) with more constraints. If some of the uncertainty sets are continuous, then (2) has an infinite number of constraints.
- It is without loss of generality to assume that the objective function has no uncertainty. We can always introduce a new variable t and minimize t subject to an additional constraint $f(x, u_0) \leq t, \forall u_0 \in \mathcal{U}_0$.

- It is not at all clear whether (2) is efficiently solvable. The computational complexity of (2) will depend on the functions $f, g_i : \mathbb{R}^n \to \mathbb{R}$ and the uncertainty sets $\mathcal{U}_i$. In general, the robust counterpart to a general convex optimization problem is intractable. For some special classes of $f, g_i : \mathbb{R}^n \to \mathbb{R}$ and special uncertainty sets $\mathcal{U}_i$, there are tractable reformulations.

In this lecture, **under some assumptions**, we introduce the following hierarchies

Robust LP	→	SOCP
Robust SOCP	→	SDP
Robust SDP	→	SOS

Remark 17.1. *Robustness is a natural requirement in many science and engineering problems. Another very related field is robust control. Indeed, much of the motivation of developing robust optimization came from the robust control community [4]. We have seen a few other important results in optimization (e.g., S-lemma, S-procedure, early developments/applications of SDPs/SOS) are from control communities. Classical text books on robust control can refer to [5, 8].* □

2 Robust Linear programs

In this section, we consider the robust counterpart of linear programs, and we consider two types of uncertainty sets: polytopic and ellipsoidal. A robust LP is a problem of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & a_i^\mathsf{T} x \le b_i, \qquad \forall a_i \in \mathcal{U}_i, i = 1, \ldots, m. \end{aligned} \tag{3}$$

2.1 Robust LP with polytopic uncertainty

Here, we assume the uncertainty set $\mathcal{U}_i$ is a polyhedron, i.e.,

$$\mathcal{U}_i = \{a_i \in \mathbb{R}^n \mid D_i a_i \le d_i\},$$

where $D_i \in \mathbb{R}^{k_i \times n}, d_i \in \mathbb{R}^{k_i}$ are given problem data. It is clear that (3) can be equivalently written as

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & \max_{a_i \in \mathcal{U}_i} a_i^\mathsf{T} x \le b_i, \qquad i = 1, \ldots, m. \end{aligned} \tag{4}$$

This a min-max problem. The strategy is to change this min-max problem to a min-min problem and then combine two minimization problems into a single one. To do this, we first deal with the inner maximization problem in (4), which is $i = 1, \ldots, m$

$$\begin{aligned} \max_{a_i \in \mathbb{R}^n} \quad & a_i^\mathsf{T} x \\ \text{subject to} \quad & D_i a_i \le d_i. \end{aligned} \tag{5}$$

By strong duality of LPs ($\mathcal{U}_i$ is naturally feasible), we can obtain the same objective value of (5) by looking at the dual of (5), which is given as

$$\begin{aligned} \min_{p_i \in \mathbb{R}^{k_i}} \quad & d_i^\mathsf{T} p_i \\ \text{subject to} \quad & D_i^\mathsf{T} p_i = x, \\ & p_i \ge 0. \end{aligned} \tag{6}$$

Therefore, we can rewrite (4) into

$$\begin{array}{ll} \min\limits_{x} & c^\top x \\ \text{subject to} & \begin{bmatrix} \min\limits_{p_i \in \mathbb{R}^{k_i}} & d_i^\top p_i \\ \text{subject to} & D_i^\top p_i = x, \\ & p_i \geq 0. \end{bmatrix} \leq b_i, \qquad i = 1, \ldots, m. \end{array} \tag{7}$$

This min-min problem (7) is also equivalent to

$$\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n, p_i \in \mathbb{R}^{k_i}} & c^\top x \\ \text{subject to} & d_i^\top p_i \leq b_i, \quad i = 1, \ldots, m \\ & D_i^\top p_i = x, \quad i = 1, \ldots, m \\ & p_i \geq 0, \quad i = 1, \ldots, m. \end{array} \tag{8}$$

The equivalence is not difficult to see:

- Suppose we have an optimal solution x, p_i for (8). Then x is also feasible for (7) with the same objective value.
- Suppose x is feasible for (7). There must exist p_i verifying the inner LP constraint. Now, (x, p_i) is also feasible to (8) and gives the same objective value.

In summary, the strong duality of LPs allows us to solve a robust LP with polytopic uncertainty (4) by solving a regular LP (of a larger size) in (8).

2.2 Robust LP with ellipsoidal uncertainty

In this section, we consider the robust LP (3) with ellipsoidal uncertainty

$$\mathcal{U}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}, i = 1, \ldots, m,$$

where $P_i \in \mathbb{R}^{n \times n}, \bar{a}_i \in \mathbb{R}^n$ are problem data. It is clear that the sets $\mathcal{U}_i$ are ellipsoids, which justify the name ellipsoidal uncertainty. If $P_i = I$, then the uncertainty sets are exactly spheres. If $P_i = 0$, there is no uncertainty. We have seen this problem in Lecture 5.

Again, we can formulate the same min-max problem in (4). In this case, the inner maximization problem has an explicit solution. Indeed, we have

$$\max_{a_i \in \mathcal{U}_i} a_i^\top x = \bar{a}_i^\top x + \max_{\|u\|_2 \leq 1} u^\top P_i^\top x = \bar{a}_i^\top x + \|P_i^\top x\|_2,$$

where the last step is from the fact that the dual norm of l_2 norm is l_2 norm. Thus, (4) can be rewritten as

$$\begin{array}{ll} \min\limits_{x} & c^\top x \\ \text{subject to} & \bar{a}_i^\top x + \|P_i^\top x\|_2 \leq b_i, \qquad i = 1, \ldots, m, \end{array} \tag{9}$$

which is an SOCP.

In summary, a robust LP with ellipsoidal uncertainty can be solved efficiently by solving a single SOCP (9).

3 Robust SOCP and (convex) QCQP

Convex QCQPs in (2) have functions g_i of the form

$$g_i(x, u_i) = \|A_i x\|_2^2 + b_i^\top x + c_i, \quad i = 1, \ldots, m.$$

SOCPs in (2) have functions of the form

$$g_i(x, u_i) = \|A_i x + b_i\| - c_i^\top x - d_i, \quad i = 1, \ldots, m.$$

In both cases, if the uncertainty $\mathcal{U}$ is a single ellipsoid (called simple ellipsoidal uncertainty), the robust counterpart is amount to solve an SDP. If $\mathcal{U}$ is polyhedral or the intersection of ellipsoids, the robust counterpart is NP-hard [3].

Here, we only briefly describe how to obtain an explicit reformulation of a robust quadratic constraint, subject o a single ellipsoidal uncertainty (see [3] for details). We consider a quadratic constraint

$$x^\top A^\top A x \le 2b^\top x + c, \qquad \forall (A, b, c) \in \mathcal{U}, \tag{10}$$

where the uncertainty set $\mathcal{U}$ is an ellipsoid about a normal point (A_0, b_0, c_0),

$$\mathcal{U} = \left\{ (A, b, c) = (A_0, b_0, c_0) + \sum_{i=1}^{l} u_i (A_i, b_i, c_i) \,\middle|\, \|u\|_2 \le 1 \right\},$$

with $(A_i, b_i, c_i), i = 0, 1, \ldots, l$ being problem data.

It is clear that a point x satisfies (10) if and only if $p^* \le 0$ where

$$\begin{aligned} p^* := \max_{A,b,c} \quad & x^\top A^\top A x - 2b^\top x - c \\ \text{subject to} \quad & (A, b, c) \in \mathcal{U} \end{aligned} \tag{11}$$

Problem (11) is the maximization of a convex quadratic objective subject to a single quadratic constraint. This problem is non-convex, but we can apply S-lemma to transform it into an SDP and solve it exactly (see S-lemma in Lecture 11).

Taking the dual of the SDP resulting from the S-lemma, we have an exact SDP for the inner maximization problem in the robust optimization problem. In particular, x is feasible to (10) if and only if there exists a scalar $\tau \in \mathbb{R}$ such that the following linear matrix inequality is feasible

$$\begin{bmatrix} c_0 + 2x^\top b_0 - \tau & \frac{1}{2}c_1 + x^\top b_1 & \ldots & \frac{1}{2}c_l + x^\top b_l & (A_0 x)^\top \\ \frac{1}{2}c_1 + x^\top b_1 & \tau & & & (A_1 x)^\top \\ \vdots & & \ddots & & \vdots \\ \frac{1}{2}c_l + x^\top b_l & & & \tau & (A_l x)^\top \\ A_0 x & A_1 x & \ldots & A_l x & I \end{bmatrix} \succeq 0$$

In summary, the robust counterpart of convex QCQPs and SOCPs subject to a single ellipsoid uncertainty is amount to solve a single SDP [3].

4 Robust semidefinite programs

Finally, we briefly describe a robust version of SDPs; see [6, 7] for more details. Consider a robust SDP of the form

$$\begin{aligned} \inf_{\lambda \in \mathbb{R}^\ell} \quad & b^\top \lambda \\ \text{subject to} \quad & P(x, \lambda) := P_0(x) - \sum_{i=1}^{\ell} P_i(x)\lambda_i \succeq 0 \quad \forall x \in \mathcal{K}, \end{aligned} \tag{12}$$

where $\lambda \in \mathbb{R}^\ell$ is the optimization variable, $P_0, \ldots, P_\ell$ are $m \times m$ symmetric polynomial matrices depending on the uncertainty parameter $x \in \mathbb{R}^n$, and the uncertain set

$$\mathcal{K} = \{x \in \mathbb{R}^n : \ g_1(x) \geq 0, \ldots, g_q(x) \geq 0\} \tag{13}$$

is a basic semialgebraic set defined by inequalities on fixed polynomials $g_1, \ldots, g_q$.

Verifying polynomial matrix inequalities is generally an NP-hard problem, which makes (12) intractable. Nevertheless, feasible vectors λ can be found via semidefinite programming if one imposes the stronger SOS condition that

$$P(x, \lambda) = S_0(x) + g_1(x)S_1(x) + \cdots + g_q(x)S_q(x) \tag{14}$$

for some $m \times m$ sum-of-squares (SOS) polynomial matrices $S_0, \ldots, S_q$. A polynomial matrix $S(x)$ is SOS if $S(x) = H(x)^{\mathsf{T}} H(x)$ for some polynomial matrix $H(x)$, and it is well known that linear optimization problems with SOS matrix variables can be reformulated as semidefinite programs (SDPs).

Now, we can solve an SOS program to get an upper bound for (12), which is

$$\begin{aligned} \inf_{\lambda \in \mathbb{R}^\ell} \quad & b^{\mathsf{T}} \lambda \\ \text{subject to} \quad & P(x, \lambda) := S_0(x) + g_1(x)S_1(x) + \cdots + g_q(x)S_q(x), \\ & S_1(x), \ldots, S_q(x) \ \text{ are SOS matrices.} \end{aligned} \tag{15}$$

Upon fixing the degrees of $S_1(x), \ldots, S_q(x)$, the problem (15) can be reformulated into a larger SDP. Under mild assumptions (the set $\mathcal{K}$ is Archimedean and there exists a λ_0 such that $P(x, \lambda_0)$ is strictly positive definite on $\mathcal{K}$), the optimal cost value of (15) will converge to the optimal cost value of (12) when we increase the degree of the SOS matrices $S_1(x), \ldots, S_q(x)$; see [6, Theorem 2] and [7, Theorems 3.1 & 3.2].

In summary, if the set $\mathcal{K}$ is Archimedean and there exists a λ_0 such that $P(x, \lambda_0)$ is strictly positive definite on $\mathcal{K}$, then the robust SDP (12) is amount to solve a sequence of SOS programs of larger sizes (15) [6, Theorem 2] and [7, Theorems 3.1 & 3.2].

Notes

The preparation of this lecture was based on [1, Lecture 16] and [4]. Further reading for this lecture can refer to [2, 4].

References

[1] Amir Ali Ahmadi. ORF523: Convex and Conic Optimization, Spring 2021.

[2] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton university press, 2009.

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

[4] Dimitris Bertsimas, David B Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

[5] Geir E Dullerud and Fernando Paganini. *A course in robust control theory: a convex approach*, volume 36. Springer Science & Business Media, 2013.

[6] Carsten W Scherer and Camile WJ Hol. Matrix sum-of-squares relaxations for robust semi-definite programs. *Mathematical programming*, 107(1):189–211, 2006.

[7] Yang Zheng and Giovanni Fantuzzi. Sum-of-squares chordal decomposition of polynomial matrix inequalities. *Mathematical Programming*, pages 1–38, 2021.

[8] Kemin Zhou, John C Doyle, and Keith Glover. Robust and optimal control, 1996.