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

Lecture 6: LP, QP, QCQP, SOCP, and SDP

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. Linear programming (LP)
2. Convex quadratic programming (QP)
3. Convex Quadratically Constrained Quadratic Programming (QCQP)
4. Second-order Cone Programming (SOCP)
5. Semidefinite Programming (SDP)

In this lecture, we will cover some of the most well-known classes of convex optimization problems and their applications, including LP, QP, QCQP, SOCP, and SDP. As we will see, SDP is the most general class of convex optimization problems among them.

1 Convex functions and Convex Optimization problems

Domain of a function: The domain of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the set $\mathrm{dom} f \subseteq \mathbb{R}^n$ where the function f is well-defined, i.e.,

$$\mathrm{dom} f := \{x \in \mathbb{R}^n \mid -\infty < f(x) < +\infty\}.$$

For example, the function $f(x) = \log(x)$ has domain $\mathrm{dom} f = \mathbb{R}_{++}$, and the function $f(X) = \log\det(X)$ has domain $\mathrm{dom} f = \mathbb{S}^n_{++}$.

Definition 6.1 (Convex function)**.** *A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain $dom f$ is a convex set and $\forall x, y \in dom f, \lambda \in [0, 1]$, we have*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

A function f is called concave if $-f$ is convex. Here are some examples

- The affine function $f(x) = c^\mathsf{T} x + b$ is convex (concave);
- The indicator function of a set

  $$\mathbb{I}_C(x) = \begin{cases} 0, & x \in C \\ \infty, & x \notin C \end{cases}$$

  is convex if and only if the set is convex
- The quadratic function $f(x) = x^\mathsf{T} Px + 2q^\mathsf{T} x + r$ with $P \in \mathbb{S}^n_+$ is convex;

- The function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) = \begin{cases} \frac{1}{x}, & x > 0 \\ \infty, & x \leq 0 \end{cases}$$

  is convex.

- The function $f(x) = \log(x)$ is concave.

There are many alternative ways to characterize the convexity of a function. Perhaps, the most commonly known one is: if $f(x)$ is twice differentiable, then it is convex if and only if its domain $\mathrm{dom} f$ is convex and its Hessian is positive semidefinite,

$$\nabla^2 f(x) \succeq 0, \ \forall x \in \mathrm{dom} f.$$

Convex optimization problems: A convex optimization problem is a problem of the form

$$\begin{array}{rl} \min & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_j(x) = 0, \quad j = 1, \ldots, p, \end{array} \tag{1}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}, i = 0, 1, \ldots, m$ are convex functions and $h_i : \mathbb{R}^n \to \mathbb{R}$ are affine functions.

We denote the feasible set as

$$\Omega = \{x \in \mathbb{R}^n \mid f_i(x) \leq 0, i = 1, \ldots, m; h_j(x) = 0, j = 1, \ldots, p\}.$$

The feasible set Ω is always convex.

- **Optimal solution**: An (globally) optimal solution x^* is a point in Ω such that

$$f_0(x^*) \leq f(x), \forall x \in \Omega.$$

  Note note an optimal solution may not exist, or may not be unique.

- **Optimal value:** The optimal value p^* of (1) is defined as

$$p^* = \inf_{x \in \Omega} f_0(x).$$

  If the problem is *infeasible*, i.e., $\Omega = \emptyset$, then $p^* = \infty$. If there exist a sequence of feasible points $x_k \in \Omega$ such that

$$f(x_k) \to -\infty,$$

  as $k \to \infty$, then $p^* = -\infty$, and we say the problem is *unbounded below*.

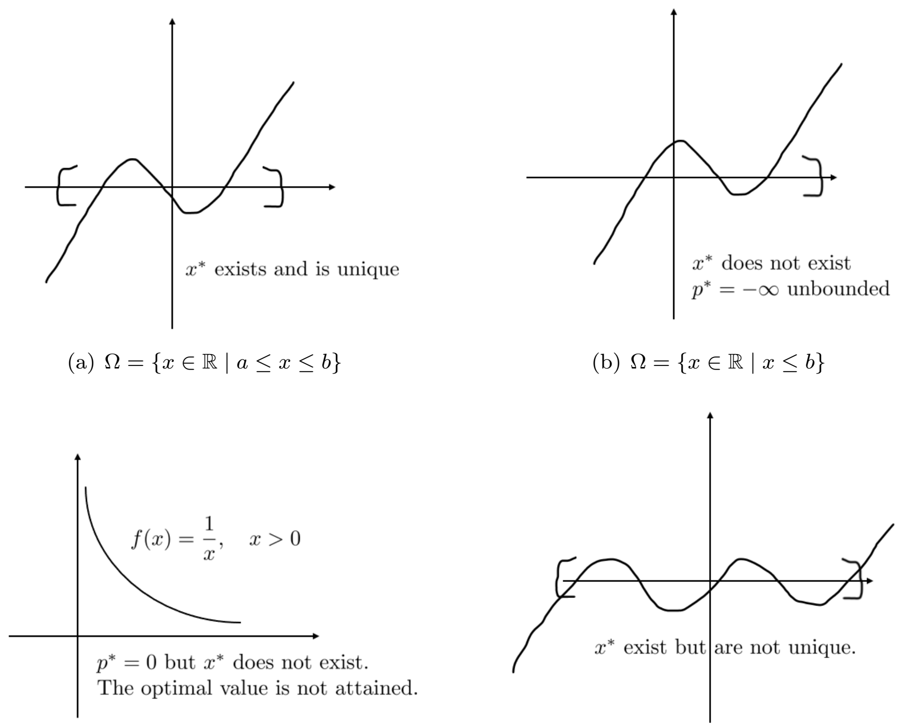
- **Solvable problem:** if there exist a point x such that $f(x) = p^*$, it is an optimal point and we say the optimal value is attained or achieved. The problem is *solvable*.

Figure 1 illustrates the concepts above.

2 A hierarchy of convex optimization problems

2.1 Linear programming

A linear program (LP) is an optimization problem of the form (1), where every function $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ are affine. The feasible region of an LP is a polyhedron.



(a) $\Omega = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ (b) $\Omega = \{x \in \mathbb{R} \mid x \leq b\}$

(c) $\Omega = \{x \in \mathbb{R} \mid x \geq 0\}$ (d) $\Omega = \{x \in \mathbb{R} \mid a \leq x \leq b\}$

Figure 1: Different situations of an optimal solutions: (a) the optimal solution is unique; (b) the optimal solution does not exist and the problem is unbounded; (c) the optimal cost is finite but it is not attained; (d) the optimal solutions are not unique.

In Lecture 4, we have introduced the standard form of an LP

$$\begin{aligned} \min_{x} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned} \tag{2}$$

l_1 **norm regression:** The problem has the form

$$\min_{x \in \mathbb{R}^n} \quad \sum_{i=1}^{m} |a_i^\mathsf{T} x + b_i|,$$

which is not an LP in its present form. As we see in Lecture 4, it can be transformed into an equivalent LP by introducing additional variables

$$\begin{aligned} \min_{x \in \mathbb{R}^n, v \in \mathbb{R}^m} \quad & \sum_{i=1}^{m} v_i \\ \text{subject to} \quad & a_i^\mathsf{T} x + b_i \geq -v_i, \\ & a_i^\mathsf{T} x + b_i \leq v_i, i = 1, \ldots, m. \end{aligned}$$

Piece-wise linear minimization. A piece-wise linear function has the form

$$f(x) = \max_{1 \leq i \leq m} a_i^\mathsf{T} x + b_i.$$

The problem of minimizing the piece-wise linear function is not an LP since the function $f(x)$ is convex but not affine. However, it is equivalent to an LP as follows

$$\begin{aligned} \min_{x\in\mathbb{R}^n, t} \quad & t \\ \text{subject to} \quad & a_i^\mathsf{T} x + b_i \leq t, i = 1, \ldots, m. \end{aligned}$$

2.2 Quadratic programming

A quadratic program is an optimization problem of the form

$$\begin{aligned} \min_{x} \quad & x^\mathsf{T} Q x + q^\mathsf{T} x + c \\ \text{subject to} \quad & Ax = b, \\ & Hx \leq d, \end{aligned} \tag{3}$$

where $Q \in \mathbb{S}^n, q \in \mathbb{R}^n, c \in \mathbb{R}, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, H \in \mathbb{R}^{q\times n}, d \in \mathbb{R}^q$.

When $Q \succeq 0$, this problem is called convex quadratic programming (QP); otherwise, it is called nonconvex QP. The case of convex QP can be solved easily, while the case of nonconvex QP is very hard to solve.

$$LP \subset QP \;\; (\text{take } Q = 0)$$

Model Predictive Control: QP has many practical applications. We here describe a useful application in model predictive control (MPC). Consider a discrete-time dynamical system

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t, \\ y_t &= Cx_t, \end{aligned}$$

where $x_t \in \mathbb{R}^n, y_t \in \mathbb{R}^p, u_t \in \mathbb{R}^m, w_t \in \mathbb{R}^n$ are the system state, output, input, and disturbance at time t. Our control goal is to design the input u_t such that system outputs follow a desired trajectory $(u_{r,t}, y_{r,t})$. At each time step, an MPC controller solves the following optimization problem over the predictive horizon

$$\begin{aligned} \min_{u_t, u_{t+1}, \ldots, u_{t+N-1}} \quad & \sum_{k=1}^{N} \left(\|y_{t+k} - y_{r,t+k}\|_Q^2 + \|u_{t+k} - u_{r,t+k}\|_R^2 \right) \\ \text{subject to} \quad & x_{k+1} = Ax_k + Bu_k + w_k, \qquad k = t, t+1, \ldots, t+N-1, \\ & y_k = Cx_k, \qquad k = t, t+1, \ldots, t+N, \\ & Dy_k + Fu_k \leq d, \qquad k = t, t+1, \ldots, t+N-1, \end{aligned} \tag{4}$$

where $(u_{r,t+k}, y_{r,t+k}), k = 1, \ldots, N-1$ denote the reference trajectory in the predictive horizon, and $D \in \mathbb{R}^{l\times p}, F \in \mathbb{R}^{l\times m}, d \in \mathbb{R}^l$ define the safety constraints on the state and inputs, and $N \in \mathbb{R}$ denotes the length of the predictive horizon. It is not difficult to see that (4) is actually a QP (note that the cost function is quadratic and the constraints are linear in u_t).

2.3 Quadratically Constrained Quadratic Programming

A quadratically constrained quadratic programming (QCQP) corresponds to a problem of the form

$$\begin{aligned} \min_{x} \quad & x^\mathsf{T} Q x + q^\mathsf{T} x + c \\ \text{subject to} \quad & x^\mathsf{T} Q_i x + q_i^\mathsf{T} x + c_i \leq 0, i = 1, \ldots, m \end{aligned} \tag{5}$$

where $Q, Q_i \in \mathbb{S}^n, q, q_i \in \mathbb{R}^n, c, c_i \in \mathbb{R}$.

If all Q, Q_i are positive semidefinite, we refer to (5) as convex QCQP, which can be solved reliably. Otherwise, it is a non-convex QCQP, which can model many computationally hard problems. We will see some concrete applications later in this course.

$$QP \subset QCQP \quad (\text{take } Q_i = 0)$$

2.4 Second-order Cone Programming

A standard second-order cone problem (SOCP) is a problem of the form

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subject to} \quad & \|A_i x + b_i\| \leq c_i^\top x + d_i, i = 1, \dots, m \end{aligned} \tag{6}$$

where $A_i \in \mathbb{R}^{k_i \times n}, b_i \in \mathbb{R}^{k_i}, c_i \in \mathbb{R}^n, d_i \in \mathbb{R}$. The name comes from the fact that the variables

$$(A_i x + b_i, c_i^\top x + d_i)$$

belong to a second-order cone. In particular, a second-order cone of dimension $n+1$ is defined as

$$\mathcal{L}^{n+1} := \{(x,t) \in \mathbb{R}^{n+1} \mid \|x\|_2 \leq t\}.$$

Now, it is easy to see that (6) is equivalent to

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subject to} \quad & (A_i x + b_i, c_i^\top x + d_i) \in \mathcal{L}^{n+1}, i = 1, \dots, m. \end{aligned} \tag{7}$$

It is clear that SOCPs contain LPs as special cases by setting $A_i = 0$. Both convex QPs and convex QCQPs are special cases of SOCPs too. Too see this, we introduce a variation on the second-order cone, called the rotated second-order cone

$$\mathcal{L}_{\text{rot}}^{n+2} := \{(x,y,z) \in \mathbb{R}^{n+2} \mid 2yz \geq \|x\|_2^2, y \geq 0, z \geq 0\}.$$

It can be expressed as a linear transformation (a rotation) of the ordinary second-order cone in $\mathbb{R}^{n+2}$. Indeed, it is not difficult to convert the rotated second-order cone into the ordinary second-order cone, since the constraints above are equivalent to

$$y + z \geq \left\| \begin{bmatrix} y - z \\ \sqrt{2}x \end{bmatrix} \right\|_2 .$$

We can view QP (3) as a special case of SOCP: first, we write the problem (3) as

$$\begin{aligned} \min_x \quad & t + q^\top x + c \\ \text{subject to} \quad & Ax = b, \\ & Hx \leq d, \\ & t \geq x^\top Q x. \end{aligned} \tag{8}$$

The last constraint $t \geq x^\top Q x$ is equivalent to $(Q^{1/2}x, t, \frac{1}{2}) \in \mathcal{L}_{\text{rot}}^{n+2}$.

Similarly, we can express the constraints in convex QCQP (5) using rotated second-order cone. Thus, we can express QCQP as special cases of SOCP. On the other hand, SOCPs cannot in general be cast as QCQPs.

LP $\subset$ (convex) QP $\subset$ (convex) QCQP $\subset$ SOCP

Robust half-space constraint:

Consider a linear constraint of the form $a^\mathsf{T} x - b \leq 0$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ are problem data. In some cases, we cannot know exactly the value of a. We assume that a is only known to belong to an ellipsoid $\mathcal{E} = \{\hat{a} + Ru \mid \|u\|_2 \leq 1\}$, with center $\hat{a} \in \mathbb{R}^n$ and $R \in \mathbb{R}^{n\times k}$ given. How can we guarantee the robust version of the linear constraint

$$a^\mathsf{T} x - b \leq 0, \quad \forall a \in \mathcal{E}. \tag{9}$$

This is equivalent to impose

$$b \geq \max_{a\in\mathcal{E}} a^\mathsf{T} x = \max_{\|u\|_2\leq 1} \hat{a}^\mathsf{T} x + x^\mathsf{T} Ru = \hat{a}^\mathsf{T} x + \max_{\|u\|_2\leq 1} x^\mathsf{T} Ru = \hat{a}^\mathsf{T} x + \|R^\mathsf{T} x\|_2,$$

where the last identity applied the definition of dual norms. Therefore, a robust half-space constraint (9) is equivalent to the following SOCP constraint

$$\|R^\mathsf{T} x\|_2 \leq b - \hat{a}^\mathsf{T} x.$$

Robust linear programming:

Consider a linear program of the form

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x \\ \text{subject to} \quad & a_i^\mathsf{T} x \leq b_i, i = 1, \ldots, m. \end{aligned} \tag{10}$$

In practice, the coefficient vectors a_i may not be known perfectly, as they are subject to noise. Assume we know that

$$a_i \in \mathcal{E}_i = \{\hat{a}_i + R_i u \mid \|u\|_2 \leq 1\}.$$

In robust optimization, we aim to minimize the original objective and impose that each constraint must be satisfied for any choice of $a_i \in \mathcal{E}_i, i = 1, \ldots, m$. From the results above, we obtain a second order cone program

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x \\ \text{subject to} \quad & \hat{a}_i^\mathsf{T} x + \|R_i^\mathsf{T} x\|_2 \leq b_i, i = 1, \ldots, m. \end{aligned} \tag{11}$$

2.5 Semidefinite Programming

Semidefinite programming is the broadest class of convex optimization problems we consider in this lecture. As we introduced in the previous lecture, its standard form is

$$\begin{aligned} \min_X \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i, X\rangle = b_i, i = 1, \ldots, m \\ & X \in \mathbb{S}^n_+, \end{aligned} \tag{12}$$

where $C \in \mathbb{S}^n$ and $A_i \in \mathbb{S}^n, i = 1, \ldots, m$, $b_i \in \mathbb{R}, i = 1, \ldots, m$ are problem data.

Note that we can write an SDP (13) as an infinite LP by replacing $X \succeq 0$ with infinitely many linear constraints

$$y^\mathsf{T} X y \geq 0, \quad \forall y \in \mathbb{R}^n.$$

Alternatively, we can also write an SDP as a nonlinear program by replacing $X \succeq 0$ with $2^n - 1$ minor inequalities from Sylvester's criterion. However, it is often much more convenient in theory and in computation to treat the matrix constraint $X \succeq 0$ directly.

Attainment of optimal solutions: Even when the optimal value of SDP (13) is finite, it may not always be attained. For example

$$\begin{aligned} \min_{x_2} \quad & x_2 \\ \text{subject to} \quad & \begin{bmatrix} x_1 & 1 \\ 1 & x_2 \end{bmatrix} \succeq 0. \end{aligned}$$

Its optimal value is 0, but it is not attained.

LP as a special case of SDP. Consider the LP (2). For a vector v, we denote $\operatorname{diag}(v)$ as the diagonal matrix with v on its diagonal entries. Then, we can equivalently write LP (2) into an SDP

$$\begin{aligned} \min_{X} \quad & \langle \operatorname{diag}(c), X \rangle \\ \text{subject to} \quad & \langle \operatorname{diag}(a_i), X \rangle = b_i, i = 1, \ldots, m \\ & X \in \mathbb{S}^n_+. \end{aligned} \tag{13}$$

- LP is a special case of SDP where all matrices are extremely sparse — diagonal matrices. Note that positive semidefiniteness of a diagonal matrix is the same as nonnegativity of its diagonal elements.

- Both polyhedra (feasible region of LPs) and spectrahedra (feasible region of SDPs) are convex. The geometry of spectrahedra is far more complex than polyhedra. For example, spectrahedra can have an infinite number of extreme points. This is one fundamental reason why SDP is not naturally amenable to simplex-type algorithms.

SOCP as a special case of SDP. We first present the following fact (from the Schur complement)

$$\|x\|_2 \leq t \Longleftrightarrow \begin{bmatrix} t & x_1 & \ldots & x_n \\ x_1 & t & \ldots & 0 \\ \vdots & \ddots & & \\ x_n & 0 & \ldots & t \end{bmatrix} \succeq 0.$$

Thus, a second-order cone constraint can be written as an LMI with an "arrow" pattern.

It is clear now that SOCP (6) can be written as an SDP

$$\begin{aligned} \min_{x} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & \begin{bmatrix} c_i^\mathsf{T} x + d_i & A_i x + b_i \\ (A_i x + b_i)^\mathsf{T} & (c_i^\mathsf{T} x + d_i) I_{n-1} \end{bmatrix} \succeq 0, i = 1, \ldots, m. \end{aligned} \tag{14}$$

3 Summary

The aforementioned convex problem classes can all be written into a general conic program: Let $K \in \mathbb{R}^n$ be a proper cone. A conic program over K is an optimization problem of the form:

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subject to} \quad & Ax = b \\ & x \in K, \end{aligned} \tag{15}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are problem data.

In particular, we have seen a hierarchy of convex optimization problems

$$\text{LP} \subset \text{(convex) QP} \subset \text{(convex) QCQP} \subset \text{SOCP} \subset \text{SDP}$$

Notes

The preparation of this lecture is based on [1, Lecture 9]. Further reading for this lecture can refer to [3, Chapter 2], [4, Chapter 4], and [2, Chapters 1 & 2].

References

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

[2] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[3] Grigoriy Blekherman, Pablo A Parrilo, and Rekha R Thomas. *Semidefinite optimization and convex algebraic geometry*. SIAM, 2012.

[4] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.