1 Outline

In this lecture, we study

- Conic programming,
- Conic duality,
- Second-order cone programming.
- Convex optimization hierarchy.

2 Conic programming

Recall that a linear program (LP) is an optimization problem with a linear objective and a system of linear inequality constraints, as follows.

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax \geq b. \end{aligned} \tag{LP}$$

Here, if the rows of A are $a_1^\top, \ldots, a_n^\top$ and the components of b are $b_1, \ldots, b_n$, then the linear system $Ax \geq b$ consists of linear inequality constraints $a_1^\top x \geq b_1, \ldots, a_n^\top x \geq b_n$. Note that Ax itself is a column vector whose components are $a_1^\top x, \ldots, a_n^\top x$. Basically, the arithmetic "$\geq$" compares two column vectors Ax and b coordinatewise.

$Ax \geq b$ is equivalent to $Ax - b \geq 0$, which means that each component of the column vector $Ax - b$ is nonnegative. We know that $\mathbb{R}^n_+$ is the nonnegative orthant, that is, the set of vectors all of whose coordinates are nonnegative. Hence, $Ax - b \geq 0$ is equivalent to $\mathbb{R}^n_+$. Then the following is an equivalent expression for the above linear program.

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax - b \in \mathbb{R}^n_+. \end{aligned}$$

Let us take a closer look at the nonnegative orthant $\mathbb{R}^n_+$. It satisfies the following properties.

1. $\mathbb{R}^n_+$ is a convex cone.
2. $\mathbb{R}^n_+$ is *pointed*, which means that if $v \in \mathbb{R}^n_+$ and $-v \in \mathbb{R}^n_+$, then it must be that $v = 0$.

In fact, $\mathbb{R}^n_+$ is not just a pointed convex cone. There are other important properties of $\mathbb{R}^n_+$.

3. $\mathbb{R}^n_+$ is *closed*, which means that for any convergent sequence $\{v^n\}_{n\in\mathbb{N}}$ contained in $\mathbb{R}^n_+$, its limit $\lim_{n\to\infty} v^n$ also belongs to $\mathbb{R}^n_+$.
4. $\mathbb{R}^n_+$ has a nonempty *interior*. Equivalently, $\mathbb{R}^n_+$ contains an *interior point*. A vector v is an interior point of a set K if there exists an open ball around v which is fully contained in K. Then the interior of a set K, denoted $\text{int}(K)$, is defined as the set of all its interior points. The interior of $\mathbb{R}^n_+$ is $\mathbb{R}^n_{++}$, the positive orthant.

In summary, the nonnegative orthant $\mathbb{R}^n_+$ is a pointed and closed convex cone with a nonempty interior. In fact, there are other closed convex cones that are pointed and have a nonempty interior. For example,

- The *Lorentz cone.*

  $$\{(x_1, \ldots, x_{n-1}, x_n)^\top \in \mathbb{R}^n : \ \|(x_1, \ldots, x_{n-1})^\top\|_2 \leq x_n\}.$$

  Other equivalent names include the *second-order cone*, the *ice-cream cone*, and the *ℓ_2-norm cone*. Its interior is given by

  $$\{(x_1, \ldots, x_{n-1}, x_n)^\top \in \mathbb{R}^n : \ \|(x_1, \ldots, x_{n-1})^\top\|_2 < x_n\}.$$

- The positive semidefinite cone.

  $$\{S \in \mathbb{S}^d : \ x^\top S x \geq 0 \quad \text{for all } x \in \mathbb{R}^d\}.$$

  Its interior is the positive definite cone, the set of all positive definite matrices.

A *conic program* is an optimization problem defined with a pointed and closed convex cone K with a nonempty interior, as follows.

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax - b \in K. \end{aligned} \tag{CP}$$

Again, when $K = \mathbb{R}^n_+$, the problem reduces to a linear program. As we use the arithmetic "$\geq$" to indicate that a vector belongs to $\mathbb{R}^n_+$, we use notation "$\geq_K$" to indicate that a vector belongs to cone K. Basically, $Ax - b \in K$ is equivalent to $Ax - b \geq_K 0$ and $Ax \geq_K b$.

Example 6.1. When K is the second-order cone, the conic program (CP) is referred to as a *second-order cone program*. When K is the positive semidefinite cone, (CP) is a semidefinite program.

2.1 Conic duality

We know that the dual of the linear program (LP) is given by

$$\begin{aligned} \text{maximize} \quad & b^\top y \\ \text{subject to} \quad & A^\top y = c \\ & y \geq 0. \end{aligned} \tag{dual-LP}$$

Let us see how to derive the dual! Note that for any $y \geq 0$ (or $y \in \mathbb{R}^n_+$) and system $Ax \geq b$, we have $y^\top (Ax - b) \geq 0$ because $y \geq 0$ and $Ax - b \geq 0$. Then it follows that

$$y^\top A x \geq y^\top b.$$

If y further satisfies

$$A^\top y = c,$$

then we have

$$y^\top A x = c^\top x \geq y^\top b = b^\top y.$$

In summary, if we take $x \in \mathbb{R}^d$ satisfying $Ax \geq b$ and $y \in \mathbb{R}^n$ with $y \geq 0$ and $A^\top y = c$, then $c^\top x$ is always lower bounded by $b^\top y$. Then we can try to find the best possible lower bound by maximizing the value of $b^\top y$, which is precisely what (dual-LP) does!

Following the basic idea behind obtaining the dual linear program, we may obtain and define the dual of the conic program (CP). The *dual cone* of $K \subseteq \mathbb{R}^n$ is defined as

$$K^* = \left\{ y \in \mathbb{R}^n : \; y^\top x \geq 0 \;\; \forall x \in K \right\}.$$

The dual cone of the nonnegative orthant $\mathbb{R}^d_+$ is $\mathbb{R}^d_+$ itself.

Example 6.2. The dual cone of the positive semidefinite cone $\mathbb{S}^d_+$ is given by

$$\left\{ X \in \mathbb{R}^{d \times d} : \; \mathrm{tr}(X^\top S) = \sum_{i=1}^{d} \sum_{j=1}^{d} X_{ij} S_{ij} \geq 0 \;\; \forall S \in \mathbb{S}^d_+ \right\}.$$

In fact, the positive semidefinite cone $\mathbb{S}^d_+$ is *self-dual*, meaning that its dual cone is itself.

Theorem 6.3 (See Theorem 2.3.1 in [BTN01])**.** *Let K be a pointed and closed convex cone with nonempty interior. Then its dual cone K^* is also a pointed and closed convex cone with nonempty interior. Moreover, $(K^*)^* = K$.*

Let us see how to derive and define the dual of the conic program!

(1) Take x such that $Ax - b \in K$ and $y \in K^*$. Then $y^\top (Ax - b) \geq 0$, and therefore,

$$y^\top Ax \geq y^\top b.$$

(2) If $y \in K^*$ further satisfies $A^\top y = c$, then

$$c^\top x = y^\top Ax \geq y^\top b = b^\top y.$$

(3) Then

$$\begin{aligned} \text{maximize} \quad & b^\top y \\ \text{subject to} \quad & A^\top y = c \\ & y \in K^* \end{aligned} \tag{dual-CP}$$

provides a lower bound on the value of (CP). Here, (dual-CP) is the dual conic program of (CP).

Taking the dual of a maximization problem is similar; the dual will give an upper bound on the problem.

Example 6.4. We consider the following semidefinite program.

$$\begin{aligned} \text{maximize} \quad & \sum_{\ell=1}^{m} b_\ell y_\ell \\ \text{subject to} \quad & \sum_{\ell=1}^{m} y_\ell A_\ell \preceq C \end{aligned}$$

To obtain its dual, we take a positive semidefinite matrix X. As the positive semidefinite cone $\mathbb{S}_+^d$ is self-dual, it follows that

$$\mathrm{tr}\left(X^\top \left(C - \sum_{\ell=1}^m y_\ell A_\ell\right)\right) = \mathrm{tr}(C^\top X) - \sum_{\ell=1}^m y_\ell \cdot \mathrm{tr}((A_\ell)^\top X) \geq 0.$$

If X satisfies

$$\mathrm{tr}((A_\ell)^\top X) = b_\ell \quad \text{for } \ell = 1, \ldots, m,$$

then

$$\mathrm{tr}(C^\top X) \geq \sum_{\ell=1}^m y_\ell \cdot \mathrm{tr}((A_\ell)^\top X) = \sum_{\ell=1}^m b_\ell y_\ell.$$

This means that

$$\begin{aligned} \text{minimize} \quad & \mathrm{tr}(C^\top X) \\ \text{subject to} \quad & \mathrm{tr}((A_\ell)^\top X) = b_\ell \quad \text{for } \ell = 1, \ldots, m \\ & X \succeq 0 \end{aligned}$$

provides an upper bound on the first semidefinite program.

2.2 Conic duality theorem

We have shown that the value of the conic program (CP) is lower bounded by the value of its dual, given by (dual-CP). What is striking is that the two values coincide under some conditions! For the case of linear programming, this is known as the strong LP duality theorem. To be precise, the value of (LP) and that of (dual-LP) are the same if (LP) is feasible and bounded. It turns out that we may extend this strong duality result to general conic programs.

We have already observed that the weak duality for LP extends to conic programming. Formally,

Theorem 6.5. *The optimal value of (dual-CP) is a lower bound on that of (CP).*

Before we state the strong duality theorem for conic programming, we need to define the notion of *strict feasibility*.

Definition 6.6. We say that a solution x is *strictly feasible* to (CP) if $Ax - b$ belongs to the interior of K. When that is the case, we write it as

$$Ax - b >_K 0 \quad \text{or} \quad Ax >_K 0.$$

Moreover, we say that (CP) is *strictly feasible* if it has a strictly feasible solution.

For example, x is strictly feasible to (LP) if $Ax > b$. Now we are ready to state the following duality result.

Theorem 6.7 (See Theorem 2.4.1 in [BTN01])**.** *The following statements hold for (CP) and (dual-CP).*

1. *If (CP) is strictly feasible and bounded, then (dual-CP) is solvable and the optimal values of (CP) and (dual-CP) are the same.*
2. *If (dual-CP) is strictly feasible and bounded, then (CP) is solvable and the optimal values of (CP) and (dual-CP) are the same.*

We will learn the notion of *Lagrangian duality* later in the course, and the strict feasibility condition is analogous to the *Slater condition*.

3 Second-order cone programming

A *second-order cone program (SOCP)* is an optimization problem of the following form:

$$\begin{aligned} \text{minimize} \quad & f^\top x \\ \text{subject to} \quad & \|A_i x + b_i\|_2 \leq c_i^\top x + d_i \quad \text{for } i = 1, \dots, m, \\ & Ex = g. \end{aligned} \tag{SOCP}$$

Exercise 6.8. Prove that (SOCP) is a conic program.

3.1 Example: chance-constrained linear programming

We consider a *chance-constrained program (CCP)* given as follows.

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & \mathbb{P}\left(a^\top x \leq b\right) \geq 1 - \epsilon \end{aligned} \tag{CCP}$$

where the constraint vector a is random. The problem is to find a solution x satisfying the constraint $a^\top x \leq b$ with probability at least $1 - \epsilon$. For example, we can formulate the portfolio optimization as (CCP).

$$\begin{aligned} \text{minimize} \quad & p^\top x \\ \text{subject to} \quad & \mathbb{P}\left(r^\top x \geq 1 + \alpha\right) \geq 1 - \epsilon, \\ & 1^\top x = 1 \end{aligned}$$

where p_i is the unit price of financial asset $i \in [d]$, and r_i is the random return of asset $i \in [d]$. The goal of the problem is to find a portfolio x whose return is at least $1 + \alpha$ with probability at least $1 - \epsilon$ while minimizing the unit price of the portfolio.

When the coefficient vector a in (CCP) follows the multivariate Gaussian distribution with mean $\bar{a}$ and covariance Σ, we can reformulate it as a second-order cone program. Let $u = a^\top x$. Then u is a random variable with mean $\bar{u} = \bar{a}^\top x$ and variance $\sigma^2 = x^\top \Sigma x$. Then $\mathbb{P}\left(a^\top x \leq b\right) \geq 1 - \epsilon$ can be rewritten as

$$\mathbb{P}\left(\frac{u - \bar{u}}{\sigma} \leq \frac{b - \bar{u}}{\sigma}\right) \geq 1 - \epsilon. \tag{6.1}$$

Here, $u - \bar{u}/\sigma$ follows the Gaussian distribution with mean 0 and variance 1 whose cumulative distribution function is given by Φ, i.e.

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt.$$

Then (6.1) is equivalent to

$$\Phi\left(\frac{b - \bar{u}}{\sigma}\right) \geq 1 - \epsilon$$

and therefore is equivalent to

$$\frac{b - \bar{u}}{\sigma} = \frac{b - \bar{a}^\top x}{\|\Sigma^{1/2} x\|_2} \geq \Phi^{-1}(1 - \epsilon).$$

Hence, (CCP) can be equivalently written as

$$\begin{aligned} &\text{minimize} && c^\top x \\ &\text{subject to} && \bar{a}^\top x + \Phi^{-1}(1-\epsilon)\|\Sigma^{1/2}x\|_2 \leq b. \end{aligned}$$

3.2 Example: convex quadratic programming

Recall that a convex quadratic program can be rewritten as follows.

$$\begin{aligned} &\text{minimize} && t \\ &\text{subject to} && t \geq \frac{1}{2}x^\top Qx + p^\top x. \end{aligned}$$

Here, the constraint $t \geq \frac{1}{2}x^\top Qx + p^\top x$ can be written as a second-order cone constraint. Recall that for any positive semidefinite matrix Q, there exists a matrix P such that $Q = P^\top P$. Then $t \geq \frac{1}{2}x^\top Qx + p^\top x$ is equivalent to

$$x^\top P^\top Px \leq 2t - 2p^\top x.$$

Note that the left-hand side equals $\|Px\|_2^2$ and that

$$2t - 2p^\top x = (t - p^\top x + 1/2)^2 - (t - p^\top x - 1/2)^2.$$

This implies that $x^\top P^\top Px \leq 2t - 2p^\top x$ is equivalent to

$$\|Px\|_2^2 + (t - p^\top x - 1/2)^2 \leq (t - p^\top x + 1/2)^2,$$

which is equivalent to

$$\left\| \begin{pmatrix} Px \\ t - p^\top x - 1/2 \end{pmatrix} \right\|_2 \leq t - p^\top x + 1/2.$$

Therefore, a convex quadratic program is a second-order cone program.

3.3 Reduction to semidefinite programming

In fact, (SOCP) is an instance of semidefinite programming.

Lemma 6.9. *Let $y \in \mathbb{R}^d$. Then $\|y\|_2 \leq s$ is equivalent to*

$$\begin{pmatrix} s & y^\top \\ y & s \cdot I \end{pmatrix} \succeq 0$$

Proof. ($\Leftarrow$) Note that

$$(\|y\|_2, -y^\top) \begin{pmatrix} s & y^\top \\ y & s \cdot I \end{pmatrix} \begin{pmatrix} \|y\|_2 \\ -y \end{pmatrix} = 2(s\|y\|_2^2 - y^\top y\|y\|_2) \geq 0,$$

implying in turn that $s \geq \|y\|_2$.

($\Rightarrow$) Let $u \in \mathbb{R}$ and $v \in \mathbb{R}^d$. Then

$$\begin{aligned} (u, v^\top) \begin{pmatrix} s & y^\top \\ y & s \cdot I \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} &= u^2 s + 2uy^\top v + sv^\top v \\ &\geq \|y\|_2(u^2 + \|v\|_2^2) + 2uy^\top v \\ &\geq \|y\|_2(u^2 + \|v\|_2^2) - 2|u| \cdot \|y\|_2 \cdot \|v\|_2 \\ &\geq 0 \end{aligned}$$

where the first inequality comes from the assumption that $s \geq \|y\|_2$ and the second and third inequalities are due to the Cauchy-Schwarz inequality. Therefore, the matrix is positive semidefinite as required. □

By Lemma 6.9, any second-order cone constraint can be written as an SDP constraint.

4 Convex optimization hierarchy

We have discussed

- Linear programming (LP),
- Quadratic programming (QP),
- Semidefinite programming (SDP),
- Conic programming,
- Second-order cone programming (SOCP).

The problem classes form a hierarchy described in Figure 6.1

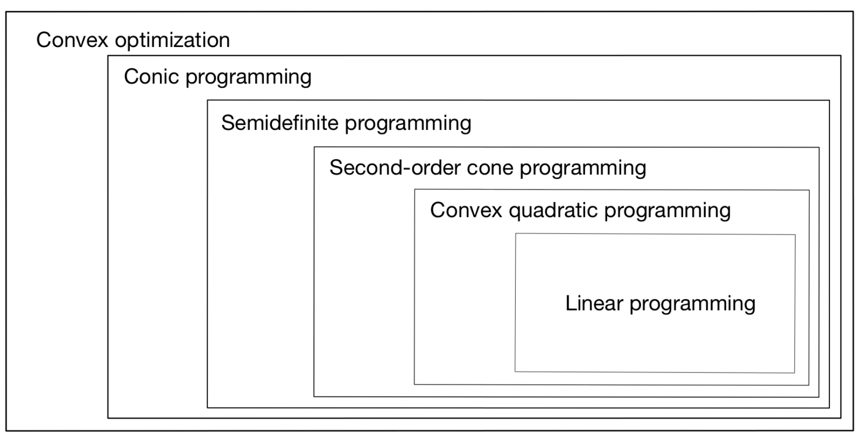

Figure 6.1: Hierarchy of classes of convex optimization problems

References

[BTN01] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization.* Society for Industrial and Applied Mathematics, 2001. 6.3, 6.7