ECE285: Semidefinite and sum-of-squares optimization	Winter 2024

Lecture 8: Duality in conic programming (II)

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. Farkas lemma
2. Strong duality
3. KKT conditions in conic programming

1 Farkas Lemma

Theorem 8.1 (Farkas' lemma)**.** *Let $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. Then, exactly one of the following sets is non-empty*

$$C = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\} \tag{1a}$$

$$D = \{y \in \mathbb{R}^m \mid A^\mathsf{T} y \leq 0, b^\mathsf{T} y > 0\} \tag{1b}$$

The system of equalities and inequalities (1a) and (1b) are called *strong alternatives*. Weak alternatives are systems where at most one of them can be feasible. The geometric interpretation of the Farkas lemma has a direct connection to the separating hyperplane theorem and makes the proof straightforward.

Geometric interpretation. We write $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$. Let $\mathrm{cone}(a_1, a_2, \dots, a_n)$ be the cone of all their nonnegative combinations. If $b \notin \mathrm{cone}(a_1, a_2, \dots, a_n)$, we can separate b from $\mathrm{cone}(a_1, a_2, \dots, a_n)$ with a hyperplane such that

$$b^\mathsf{T} y > 0, \qquad a_i^\mathsf{T} y \leq 0, \ \forall a_i \in \mathrm{cone}(a_1, a_2, \dots, a_n)$$

See Figure 1 for an illustration.

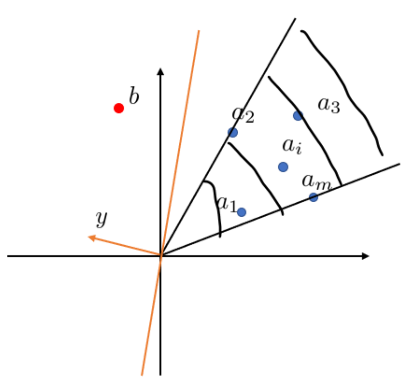

Figure 1: Geometric interpretation of the Farkas lemma

Proof of Farkas Lemma.

Proof. Suppose $C \neq \emptyset$. There exists $x \geq 0$, such that $Ax = b$. Then

$$y^\top (Ax) = y^\top b, \quad \forall y \in \mathbb{R}^m \Rightarrow \quad x^\top (A^\top y) = b^\top y, x \geq 0, \forall y \in \mathbb{R}^m.$$

If $A^\top y \leq 0$ then $b^\top y \leq 0$. Thus $D = \emptyset$.

Suppose $C = \emptyset$. We prove that $D \neq \emptyset$. We write $A = \begin{bmatrix} a_1 & a_2 & \ldots, a_n \end{bmatrix}$. Let $\text{cone}(a_1, a_2, \ldots, a_n)$ be the cone of all their nonnegative combinations. Since for any finite set of points $S = \{s_1, \ldots, s_p\}$, the set $\text{cone}(S)$ is convex and closed. Thus, $\text{cone}(a_1, a_2, \ldots, a_n)$ is convex and closed.

We can now use the separating hyperplane theorem. By assumption $C = \emptyset$, we have $b \notin \text{cone}(a_1, a_2, \ldots, a_n)$. Then, the point b and the set $\text{cone}(a_1, a_2, \ldots, a_n)$ can be strictly separated, i.e., there exist $y \in \mathbb{R}^m, y \neq 0, r \in \mathbb{R}$, such that

$$y^\top b > r, \quad y^\top z \leq r, \forall z \in \text{cone}(a_1, a_2, \ldots, a_n).$$

First $r \geq 0$ since $0 \in \text{cone}(a_1, a_2, \ldots, a_n)$. Second, Since $\text{cone}(a_1, a_2, \ldots, a_n)$ is a cone, if there exists a $z \in \text{cone}(a_1, a_2, \ldots, a_n)$ such that $y^\top z > 0$ then $y^\top (\alpha z)$ can be arbitrarily large. Therefore we can choose $r = 0$. Then, there exist $y \in \mathbb{R}^m, y \neq 0$, such that (you can also directly apply Theorem 3.10 in Lectures 3 & 4 to get the following result)

$$y^\top b > 0, \quad y^\top a_i \leq 0, \forall i = 1, \ldots n.$$

which means that $D \neq \emptyset$. □

1.1 Farkas lemma from LP strong duality

Farkas lemma can be also be directly proven from the strong duality of linear programming. The converse that the strong duality of LP can be proven from Farkas lemma is also true. Note other proofs of LP strong duality also exist, e.g., techniques based on simplex method. However, simplex-based proof does not generalize to general conic programming, while the hyperplane-based proofs are still applicable. Here, we show how to prove Farkas lemma from LP strong duality.

Consider the pair of primal and dual LPs

$$\begin{aligned} \min_{x} \quad & 0 \\ \text{subject to} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{2}$$

and

$$\begin{aligned} \max_{y,z} \quad & b^\top y \\ \text{subject to} \quad & A^\top y + z = 0 \\ & z \geq 0. \end{aligned} \tag{3}$$

Note that (3) is trivially feasible with $y = 0, z = 0$. Therefore, we have two cases

- If (2) is feasible, then by strong duality the optimal value of (3) satisfies $b^\top y \leq 0$. This means $\{y \in \mathbb{R}^m \mid A^\top y \leq 0, b^\top y > 0\}$ is infeasible.
- If (2) is infeasible, then by strong duality the optimal value of (3) is unbounded above, i.e., there exists a feasible y such that $b^\top y > 0$. This means $\{y \in \mathbb{R}^m \mid A^\top y \leq 0, b^\top y > 0\}$ is feasible.

2 Strong duality of LPs

The standard-form LP is a problem of the form

$$\begin{aligned} \min_{x} \quad & c^{\mathsf{T}} x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned} \tag{4}$$

The dual problem of (4) is

$$\begin{aligned} \max_{y,z} \quad & b^{\mathsf{T}} y \\ \text{subject to} \quad & A^{\mathsf{T}} y + z = c, \\ & z \geq 0. \end{aligned} \tag{5}$$

In Lecture 6, we have stated the following result. We will prove it using Farkas lemma.

Theorem 8.2 (Strong duality)**.** *If the primal LP* (4) *is feasible and has a finite optimal value, then the dual LP* (5) *is also feasible. There exist optimal feasible solutions* $x^* \geq 0, y^*, z^* \geq 0$ *and* $c^{\mathsf{T}} x^* = b^{\mathsf{T}} y^*$.

We first prove a variant of the Farkas lemma, which gives an infeasible certificate of a set of linear inequalities.

Lemma 8.1. *Let* $A \in \mathbb{R}^{m \times n}$. *Then,* $C = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ *is an empty set if and only if*

$$\exists y \geq 0, A^{\mathsf{T}} y = 0, b^{\mathsf{T}} y < 0.$$

Proof. $\Leftarrow$ If $\exists y \geq 0, A^{\mathsf{T}} y = 0, y^{\mathsf{T}} b < 0$, then $y^{\mathsf{T}}(Ax - b) > 0, \forall x \in \mathbb{R}^n$. Since $y \geq 0$, it means that $C = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ is empty.

$\Rightarrow$ We rewrite the LP in standard form and apply the Farkas lemma

$$\begin{aligned} & C = \{x \in \mathbb{R}^n \mid Ax \leq b\} \;\text{ empty} \\ \Leftrightarrow\; & \{(x^+, x^-, s) \mid A(x^+ - x^-) + s = b, s \geq 0, x^+ \geq 0, x^- \geq 0\} \;\text{ empty} \\ \Leftrightarrow\; & \left\{ (x^+, x^-, s) \mid \begin{bmatrix} A & -A & I \end{bmatrix} \begin{bmatrix} x^+ \\ x^- \\ s \end{bmatrix} = b, (s, x^+, x^-) \geq 0 \right\} \;\text{ empty} \\ \Leftrightarrow\; & \exists y, \text{ such that } b^{\mathsf{T}} y > 0, \begin{bmatrix} A^{\mathsf{T}} \\ -A^{\mathsf{T}} \\ I \end{bmatrix} y \leq 0 \\ \Rightarrow\; & \exists y \leq 0, \text{such that } b^{\mathsf{T}} y > 0, A^{\mathsf{T}} y = 0 \\ \Rightarrow\; & \exists y \geq 0, \text{such that } b^{\mathsf{T}} y < 0, A^{\mathsf{T}} y = 0. \end{aligned}$$

□

Proof of LP strong duality: We assume the optimal value of (4) is finite and denoted as p^*. We now aim to prove the following set of linear inequalities is feasible

$$b^{\mathsf{T}} y \geq p^*, A^{\mathsf{T}} y \leq c.$$

(Indeed, by weak duality we have $b^{\mathsf{T}} y \leq p^*$, so we get $b^{\mathsf{T}} y = p^*$).

The system is the same as

$$\begin{bmatrix} A^{\mathsf{T}} \\ -b^{\mathsf{T}} \end{bmatrix} y \leq \begin{bmatrix} c \\ -p^* \end{bmatrix}.$$

If it is infeasible, the Lemma 8.1 implies that

$$\exists \lambda := \begin{bmatrix} \lambda_1 \\ \lambda_0 \end{bmatrix} \geq 0, A\lambda_1 - b\lambda_0 = 0, c^\mathsf{T}\lambda_1 - p^*\lambda_0 < 0.$$

which means that

$$\exists \lambda_1 \geq 0, \lambda_0 \geq 0, \text{ such that } A\lambda_1 = b\lambda_0 \text{ and } c^\mathsf{T}\lambda_1 < p^*\lambda_0. \tag{6}$$

We have two cases:

- Case 1: $\lambda_0 = 0$. (6) leads to $\exists \lambda_1 \geq 0$ such that $A\lambda_1 = 0$ and $c^\mathsf{T}\lambda_1 < 0$. Suppose an optimal solution of (4) is $x^* \geq 0, c^\mathsf{T}x^* = p^*$. We let $x = x^* + \lambda_1 \geq 0$, which is feasible

  $$Ax = A(x^* + \lambda_1) = b$$

  and

  $$c^\mathsf{T}x = c^\mathsf{T}x^* + c^\mathsf{T}\lambda_1 < p^*.$$

  This contradicts the fact that p^* is the primal optimal value.

- Case 2: $\lambda_0 > 0$. We let

  $$x = \frac{\lambda_1}{\lambda_0} \geq 0,$$

  and (6) leads to

  $$Ax = b, c^\mathsf{T}x < p^*.$$

  This contradicts the fact that p^* is the primal optimal value.

3 KKT condtions in conic programming

Let $K \in \mathbb{R}^n$ be a proper cone. A conic program over K is an optimization problem of the form:

$$\begin{aligned} p^* = \min_x \quad & c^\mathsf{T}x \\ \text{subject to} \quad & Ax = b \\ & x \in K, \end{aligned} \tag{7}$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are problem data. The dual of (7) is a maximization problem of the form

$$\begin{aligned} d^* = \max_{y,z} \quad & b^\mathsf{T}y \\ \text{subject to} \quad & A^\mathsf{T}y + z = c \\ & z \in K^*. \end{aligned} \tag{8}$$

Theorem 8.3 (Duality for conic programs)**.** *Consider the conic program* (7) *and its dual* (8)*. The following statements hold*

1. *Weak duality:* $p^* \geq d^*$*;*
2. *Strong duality:* $p^* = d^*$ *if problem* (7) *is strictly feasible (i.e., there exists* $x \in int(K)$ *such that* $Ax = b$*).*

This result works for any conic programs, such as LP, SOCP, and SDPs. Note that for the pair of primal and dual LPs (4)-(5), we only require a feasible (not necessarily strictly feasible) primal solution to guarantee strong duality.

From Theorem 8.3, we can derive the following KKT (Karush-Kuhn-Tucker) conditions.

Theorem 8.4 (KKT conditions)**.** *Consider the conic program* (7) *and its dual* (8)*. Suppose that* (7) *is strictly feasible and that both primal and dual optimal solutions are attained. Then a point x is optimal if and only if there exists y, z such that*

- *Primal feasibility: $x \in K$, and $Ax = b$;*
- *Dual feasibility: $z \in K^*$, and $c = z + A^\mathsf{T} y$;*
- *Complementary slackness $x^\mathsf{T} z = 0$.*

Proof $\Rightarrow$ From strong duality Theorem 8.3, we know there exist $x \in K$, $z \in K^*$, such that

$$Ax = b, c = z + A^\mathsf{T} y, c^\mathsf{T} x = b^\mathsf{T} y.$$

It follows that

$$c^\mathsf{T} x = (z + A^\mathsf{T} y)^\mathsf{T} x = z^\mathsf{T} x + y^\mathsf{T} Ax = b^\mathsf{T} y, \tag{9}$$

then $z^\mathsf{T} x = 0$.

$\Leftarrow$ Suppose the KKT conditions hold, we have

$$c^\mathsf{T} x = (z + A^\mathsf{T} y)^\mathsf{T} x = z^\mathsf{T} x + y^\mathsf{T} Ax = b^\mathsf{T} y, \quad x \in K, z \in K^* \tag{10}$$

On the other hand, we have $c^\mathsf{T} x \geq b^\mathsf{T} y$ for any feasible x, z. Thus, x is an optimal solution of (7), and

$$c^\mathsf{T} x = p^* = d^* = b^\mathsf{T} y.$$

3.1 Case of LPs

The KKT conditions for the pair of primal-dual LPs (4)-(5) take the form

- Primal feasibility: $x \geq 0$, and $Ax = b$;
- Dual feasibility: $z \geq 0$, and $c = z + A^\mathsf{T} y$;
- Complementary slackness $x_i z_i = 0, i = 1, \ldots, n$.

Complementary slackness for a general conic program is $x^\mathsf{T} z = 0$. In the case of LP, this condition is equivalent to $x_i z_i = 0, i = 1, \ldots, n$ since $x \geq 0, z \geq 0$.

3.2 Case of SDPs

The standard form of SDPs is

$$\begin{aligned} \min_X \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \ldots, m \\ & X \in \mathbb{S}^n_+, \end{aligned} \tag{11}$$

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. The dual of (11) is

$$\begin{aligned} \max_{y, Z} \quad & b^\mathsf{T} y \\ \text{subject to} \quad & \sum_{i=1}^m y_i A_i + Z = C \\ & Z \in \mathbb{S}^n_+, \end{aligned} \tag{12}$$

The KKT conditions for the pair of primal-dual SDPs (11)-(12) take the form

- Primal feasibility: $X \succeq 0$, and $\langle A_i, X\rangle = b_i, i = 1, \ldots, m$;
- Dual feasibility: $Z \succeq 0$, and $C = Z + \sum_{i=1}^m A_i^\mathsf{T} y_i$;
- Complementary slackness $XZ = 0$.

We have proved that $\text{trace}(XZ) = 0 \Leftrightarrow XZ = 0$ when $X \succeq 0, Z \succeq 0$ (Problem 1 in Homework 1).

Example 8.1 (KKT conditions in SDPs [3, Examples 2.11 & 2.13])**.** *Consider a primal-dual pair of SDPs*

$$\begin{aligned} \min_{X} \quad & 2x_{11} + 2x_{12} \\ \textit{subject to} \quad & x_{11} + x_{22} = 1, \\ & \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \succeq 0, \end{aligned}$$

and

$$\begin{aligned} \max_{y} \quad & y \\ \textit{subject to} \quad & \begin{bmatrix} 2-y & 1 \\ 1 & -y \end{bmatrix} \succeq 0. \end{aligned}$$

Both the primal and dual SDPs are strictly feasible, thus strong duality holds. Indeed, the primal optimal solution is

$$X^* = \begin{bmatrix} \frac{2-\sqrt{2}}{4} & -\frac{1}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & \frac{2+\sqrt{2}}{4} \end{bmatrix}$$

with the primal optimal cost $p^* = 1 - \sqrt{2}$. The dual optimal solution is $y^* = 1 - \sqrt{2}$ with the dual optimal cost $d^* = 1 - \sqrt{2}$. Furthermore, complementary slackness holds:

$$\left(C - \sum_{i=1}^m A_i y_i^*\right) X^* = \begin{bmatrix} 1+\sqrt{2} & 1 \\ 1 & \sqrt{2}-1 \end{bmatrix} \begin{bmatrix} \frac{2-\sqrt{2}}{4} & -\frac{1}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & \frac{2+\sqrt{2}}{4} \end{bmatrix} = 0.$$

□

Notes

The preparation of this lecture was based on [1, Lecture 5] and [5, Lecture 6]. Further reading for this lecture can refer to [3, Chapter 2], [2, Chapter 1] and [4, Chapter 4].

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.

[5] Hamza Fawzi. Topics in Convex Optimisation, Michaelmas 2018.