Chapter 4: Duality Theory

June 18, 2022

Scribe: Eric Chen

Throughout this chapter, we consider the standard form problem

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

and let P be the corresponding feasible set. We assume that the dimensions of matrix A are $m \times n$ and that its rows are linearly independent. Also, denote A_i as the ith column of the matrix A and a_i^T its ith row.

In summary, the construction of the dual of a primal minimization problem can be viewed as follows. We have a vector of parameters p, and for every p we have a method for obtaining a lower bound on the optimal cost. The dual problem is a maximization problem that looks for the tightest such lower bound.

1 The Dual Problem

Given primal problems and dual problems, we summarize their relations in the following table.

PRIMAL	minimize	maximize	DUAL
constraints	$\geq b_i$	≥ 0	**variables**
constraints	$\leq b_i$	≤ 0	**variables**
constraints	$= b_i$	free	**variables**
variables	≥ 0	$\leq c_j$	**constraints**
variables	≤ 0	$\geq c_j$	**constraints**
variables	free	$= c_j$	**constraints**

Its proof can be completed by simply following the definition.

Theorem 1.1. *If we transform the dual into an equivalent minimization problem and then form its dual, we obtain a problem equivalent to the original problem.*

Theorem 1.2. *Suppose that we have transformed a linear programming problem Π_1 to another linear programming problem Π_2, by a sequence of transformations of the following types:*

(a) Replace a free variable with the difference of two nonnegative variables.

(b) Replace an inequality constraint by an equality constraint involving a nonnegative slack variable.

(c) If some row of the matrix A in a feasible standard form problem is a combination of the other rows, eliminate the corresponding equality constraint.

Then, the duals of Π_1 and Π_2 are equivalent, i.e., they are either both infeasible or they have the same optimal cost.

2 The Duality Theorem

Theorem 2.1 (Weak duality). *If x is a feasible solution to the primal problem and p is a feasible solution to the dual problem, then*

$$p^T b \leq c^T x.$$

Proof. For any vectors x and p, we define

$$u_i = p_i(a_i^T x - b_i), \quad v_i = (c_j - p^T A_j)x_j.$$

The definition of the dual problem requires the sign of p_i to be the same as the sign of $a_i^T x - b_i$ and the sign of $c_j - p^T A_j$ to be the same as the sign of x_j. Thus we have:

$$u_i, v_j \geq 0 \;\; \forall i, j.$$

Notice that

$$\sum_i u_i = p^T A x - p^T b,$$

and

$$\sum_j v_j = c^T x - p^T A x.$$

By adding two equalities, we obtain:

$$0 \leq \sum_i u_i + \sum_j v_j = c^T x - p^T b.$$

A more simpler way to approach this proof is through Lagrangian function, as weak duality just means that

$$\sup_p \inf_x \mathcal{L}(x,p) \leq \inf_x \sup_p \mathcal{L}(x,p).$$

□

By the weak duality theorem, we have the following lemma:

Lemma 2.2.

(a) If the optimal cost in the primal is $-\infty$, then the dual problem must be infeasible.

(b) If the optimal cost in the dual is $+\infty$, then the primal problem must be infeasible.

Another lemma of the weak duality theorem is shown as follows.

Lemma 2.3. *Let x and p be feasible solutions to the primal and the dual, respectively, and suppose that $p^T b = c^T x$. Then x and p are optimal solutions to the primal and the dual, respectively.*

We should notice that linear programming problem satisfies strong duality.

Theorem 2.4 (Strong duality). *If a linear programming problem has an optimal solution, so does its dual, and the respective optimal costs are equal.*

Proof. We prove this by simplex method and assume this problem is in standard form. Let us assume temporarily that the rows of A are linearly independent and that there exists an optimal solution. Also, we use some rules to avoid cycling, thus the simplex method terminates with an optimal solution x and an optimal basis B. Let $x_B = B^{-1}b$. When it terminates, we have the reduced costs

$$c^T - c_B^T B^{-1} A \geq 0.$$

By letting $p^T = c_B^T B^{-1}$, we have $p^T A \leq c^T$, where p is thus a feasible solution to the dual problem. In addition,

$$p^T b = c_B^T B^{-1} b = c_B^T x_B = c^T x.$$

We have thus shown that the optimal dual cost is equal to the optimal primal cost. □

By the weak and strong duality theorem, we can derive the following table:

	Finite optimum	**Unbounded**	**Infeasible**
Finite optimum	Possible	Impossible	Impossible
Unbounded	Impossible	Impossible	Impossible
Infeasible	Impossible	Possible	Possible

Another important relation between primal and dual optimal solutions is provided by the *complementary slackness* conditions.

Theorem 2.5 (Complementary slackness). *Let x and p be feasible solutions to the primal and the dual problem, respectively. The vectors x and p are optimal solutions for the two respective problems if and only if:*

$$p_i(a_i^T x - b_i) = 0, \qquad \forall i,$$

and

$$(c_j - p^T A_j)x_j = 0, \qquad \forall j.$$

Proof. By the proof of weak duality, we showed that

$$c^T x - p^T b = \sum_i u_i + \sum_j v_j.$$

By the strong duality theorem, if x and p are optimal, then $c^T x = p^T b$, which implies that $u_i = v_j = 0$ for all i, j. $\square$

3 Farkas' Lemma and Linear Inequalities

Theorem 3.1 (Farkas' lemma). *Let A be a matrix of dimensions $m \times n$ and let b be a vector in $\mathbb{R}^m$. Then exactly one of the following two alternatives holds:*

(a) There exists some $x \geq 0$ such that $Ax = b$.

(b) There exists some vector p such that $p^T A \geq 0$ and $p^T b < 0$.

Proof. Suppose that (a) holds, then if $p^T A \geq 0$, then $p^T b = p^T Ax \geq 0$, which shows that $p^T b < 0$ cannot hold.
Now assume that (a) doesn't hold. Then consider

$$\begin{aligned} \min \quad & p^T b \\ \text{s.t.} \quad & p^T A \geq 0 \end{aligned}$$

and problem

$$\begin{aligned} \max \quad & 0^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Notice that the second is the dual of the first problem. The maximization problem is infeasible, which implies that the minimization problem is either unbounded or infeasible. Since $p = 0$ is a feasible solution, then the problem is unbounded. Therefore, there exists some p which is feasible, such that $p^T A \geq 0$ and $p^T b < 0$. $\square$

Below is an equivalent statement of Farkas' lemma which is sometimes more convenient.

Lemma 3.2. *Let $A_1, \ldots, A_n$ and b be given vectors and suppose that any vector p that satisfies $p^T A_i \geq 0, i = 1, \ldots, n$, must also satisfy $p^T b \geq 0$. Then b can be expressed as a nonnegative linear combinations of the vectors $A_1, \ldots, A_n$.*

Theorem 3.3. *Suppose that the system of linear inequalities $Ax \leq b$ has at least one solution, and let d be some scalar. Then the following are equivalent:*

(a) Every feasible solution to the system $Ax \leq b$ satisfies $c^T x \leq d$.

(b) There exists some $p \geq 0$ such that $p^T A = c^T$ and $p^T b \leq d$.

Proof. Consider one pair of problems, $\{\max\ c^T x \mid Ax \leq b\}$ and $\{\min\ p^T b \mid p^T A = c^T, p \geq 0\}$. Note that the first one is the dual of the second. Suppose that $c^T x \leq d$ always holds, then by the strong duality theorem, we know that there exists an optimal solution p which satisfies $p^T b \leq d$. The other side of proof can also be completed by using strong duality theorem. $\square$

4 From Separating Hyperplanes to Duality

Theorem 4.1. *Every polyhedron is closed.*

Proof. Consider $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$. As it's an intersection of finite number of closed sets, so it's closed. $\square$

Theorem 4.2 (Separating hyperplane theorem). *Suppose C and D are two convex sets that do not intersect. Then there exists $a \neq 0$ and b such that $a^T x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$.*

Proof. Define the distance between C and D as:

$$\mathbf{dist}(C, D) = \inf\{\|u - v\|_2 \mid u \in C, v \in D\},$$

and that there exists points $c \in C$, $d \in D$ that achieve the minimum distance. Define

$$a = d - c, \qquad b = \frac{\|d\|_2^2 - \|c\|_2^2}{2},$$

and define one affine function as

$$f(x) = a^T x - b = (d - c)^T (x - (1/2)(d + c)),$$

and this defines the hyperplane $\{x \mid a^T x = b\}$. First we show that f is nonpositive on D. Suppose there were a point $u \in D$ for which

$$f(u) = (d - c)^T (u - (1/2)(d + c)) < 0.$$

We can then express $f(u)$ as:

$$f(u) = (d - c)^T (u - d + (1/2)(d - c)) = (d - c)^T (u - d) + (1/2)\|d - c\|_2^2.$$

Then we know it implies that $(d - c)^T (u - d) < 0$. Now we observe that

$$\frac{d}{dt}\|d + t(u - d) - c\|_2^2 \bigg|_{t=0} = 2(d - c)^T (u - d) < 0,$$

so for some small $t > 0$ with $t \leq 1$, we have

$$\|d + t(u - d) - c\|_2 < \|d - c\|_2,$$

which means the point $d + t(u - d)$ is closer to c than d is. Since D is convex and contains d and u, we have $d + t(u - d) \in D$, which makes a contradiction. $\square$

Prove Farkas' lemma by separating hyperplanes.

Proof. Let

$$
\begin{aligned}
S &= \{Ax \mid x \geq 0\} \\
&= \{y \mid \text{there exists } x \text{ such that } y = Ax, x \geq 0\},
\end{aligned}
$$

Note that S is the projection of the polyhedron $\{(x, y) \mid y = Ax, x \geq 0\}$ onto y coordinate, is itself a polyhedron, and is closed.
We now invoke the separating hyperplane theorem to separate b from S and conclude that there exists a vector p such that $p^T b < p^T y$ for every $y \in S$. Since $0 \in S$, we must have $p^T b < 0$. Furthermore, for every column A_i of A and every $\lambda > 0$, we have $\lambda A_i \in S$ and $p^T b < \lambda p^T A_i$, then we divide both sides by λ and take λ to infinity we have $p^T A_i \geq 0$. So $p^T A \geq 0$. □

We then complete the proof of strong duality by Farkas' lemma.

Proof. Consider this pair of primal and dual problems: $\{\min\ c^T x \,|Ax \geq b\}$ and $\{\max\ p^T b \,|\, p^T A = c^T, p \geq 0\}$. We assume that the primal has an optimal solution x^*.
Let $I = \{i \mid a_i^T x^* = b_i\}$ be the active index set at x^*. Consider vector d and $\varepsilon > 0$ satisfies $a_i^T d \geq 0$. Then we have

$$a_i^T (x^* + \varepsilon d) \geq a_i^T x^* = b_i \quad \forall i \in I.$$

In addition, if $i \notin I$ and ε is sufficiently small, we know

$$a_i^T x^* > b_i \Rightarrow a_i^T (x^* + \varepsilon d) > b_i,$$

so $x^* + \varepsilon d$ is a feasible solution. By the optimality of x^*, we obtain $c^T d \geq 0$. By Farkas' lemma, c can be expressed as a nonnegative linear combination of the vectors $a_i, i \in I$:

$$c = \sum_{i \in I} p_i a_i, \ p_i > 0.$$

For $i \notin I$, we define $p_i = 0$. We then have $p \geq 0$ and $p^T A = c^T$. In addition,

$$p^T b = \sum_{i \in I} p_i b_i = \sum_{i \in I} p_i a_i^T x^* = c^T x^*,$$

which shows that the cost of this dual feasible solution p is the same as the optimal primal cost. The strong duality theorem now follows from weak duality theorem. □

5 Cones and Extreme Rays

Definition 5.1. A set $C \subset \mathbb{R}^n$ is a **cone** if $\lambda x \in C$ for all $\lambda \geq 0$ and all $x \in C$.

Remark. *A polyhedron of the form $P = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$ is a nonempty cone and is called a polyhedral cone.*

Theorem 5.2. *Let $C \subset \mathbb{R}^n$ be the polyhedral cone defined by the constraints $a_i^T x \geq 0, i = 1, \ldots, m$. Then the following are equivalent:*

(a) The zero vector is an extreme point of C.

(b) The cone C does not contain a line.

(c) There exist n vectors out of the family $a_1, \ldots, a_m$ that are linearly independent.

Remark. *If the zero vector is an extreme point, we say that the cone is* **pointed***.*

Definition 5.3. Consider a nonempty polyhedron

$$P = \{x \in \mathbb{R}^n \mid Ax \geq b\},$$

and some point $y \in P$. The **recession cone at** y is defined as the set

$$\{d \in \mathbb{R}^n \mid A(y + \lambda d) \geq b, \ \forall \lambda \geq 0\}.$$

It can be easily seen that this set is the same as

$$\{d \in \mathbb{R}^n \mid Ad \geq 0\}.$$

Remark. *From the definition above, we know that recession cone is indeed a polyhedral cone, and is independent of the starting point y. For the case of standard form*

$$P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\},$$

the recession cone is the set

$$\{d \in \mathbb{R}^n \mid Ad = 0, d \geq 0\}.$$

This can be easily verified by transforming the standard form into the corresponding general form.

Definition 5.4. The nonzero elements of the recession cone are called the **rays** of the polyhedron P.

Definition 5.5.

(a) A nonzero element x of a polyhedral cone $C \subset \mathbb{R}^n$ is called an **extreme ray** if there are $n - 1$ linearly independent constraints that are active at x.

(b) An extreme ray of the recession cone associated with a nonempty polyhedron P is also called an **extreme ray** of P.

Remark. *We say that two extreme rays are equivalent if one is a positive multiple of the other. So any $n - 1$ linearly independent constraints define a line and can lead to at most two nonequivalent extreme rays, with one being the negative of the other. Given that there is a finite number of ways that we can choose $n - 1$ constraints to become active, we conclude that the number of extreme rays of a polyhedron is finite.*

Theorem 5.6. *Consider the problem of minimizing $c^T x$ over a pointed polyhedral cone*

$$C = \{x \in \mathbb{R}^n \mid a_i^T x \geq 0, i = 1, \ldots, m\}.$$

The optimal cost is equal to $-\infty$ if and only if some extreme ray d of C satisfies $c^T d < 0$.

Proof. If there exists some extreme ray d such that $c^T d < 0$, then we can move along this direction and get $v_{opt} = -\infty$.
If the optimal cost is $-\infty$, then there must exist some $x \in C$ such that $c^T x < 0$, by suitably scaling x, we can have $c^T x = -1$. Now consider polyhedron

$$P = \{x \in \mathbb{R}^n \mid a_1^T x \geq 0, \ldots, a_m^T \geq 0, c^T x = -1\},$$

by the assumption above, we know that P is nonempty. Since C is pointed, the vectors $a_1, \ldots, a_m$ span $\mathbb{R}^n$ and implies that P has at least one extreme point, denoted as d. We know that at d there are n active linearly independent constraints, so at least $n-1$ of them are of the form $a_i^T x \geq 0$. It follows that d is an extreme ray of C. □

Theorem 5.7. *Consider the problem of minimizing $c^T x$ subject to $Ax \geq b$ and assume that the feasible set has at least one extreme point. The optimal cost is equal to $-\infty$ if and only if some extreme ray d of the feasible set satisfies $c^T d < 0$.*

Proof. If there exists some extreme ray d such that $c^T d < 0$, then the conclusion is trivial.
For the reverse direction, consider the dual problem

$$\begin{aligned} \max \quad & p^T b \\ \text{s.t.} \quad & p^T A = c^T \\ & p \geq 0 \end{aligned}$$

As the primal is unbounded, then the dual is infeasible, then the problem

$$\begin{aligned} \max \quad & p^T 0 \\ \text{s.t.} \quad & p^T A = c^T \\ & p \geq 0 \end{aligned}$$

is infeasible, which means that the associated primal problem

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq 0 \end{aligned}$$

is either unbounded or feasible. As $x = 0$ is a feasible solution, then the problem is unbounded. Since the primal feasible set has at least one extreme point, the rows of A span $\mathbb{R}^n$. It follows that the recession cone

$$\{x \mid Ax \geq 0\}$$

is pointed and thus exists an extreme ray d of the recession cone satisfying $c^T d < 0$. By definition, this is an extreme ray of the feasible set. □

6 Representation of Polyhedra

Theorem 6.1 (Resolution theorem). *Let*

$$P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$$

be a nonempty polyhedron with at least one extreme point. Let $x^1, \ldots, x^k$ be the extreme points and let $w^1, \ldots, w^r$ be a complete set of extreme rays of P. Let

$$Q = \left\{ \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j \;\middle|\; \lambda_i \geq 0, \theta_j \geq 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

Then $Q = P$.

Proof. Let

$$x = \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j$$

be an element of Q. The vector

$$y = \sum_{i=1}^{k} \lambda_i x^i$$

is a convex combination of elements of P, it therefore satisfies

$$Ay \geq b.$$

Also, let

$$z = \sum_{j=1}^{r} \theta_j w^j,$$

we know that $Az \geq 0$ as w^j are extreme rays, so $Ax = A(y+z) \geq b$. Thus we have proved $Q \subseteq P$.

For the reverse, assume that P is not a subset of Q. Let z be an element of P that does not belong to Q, consider the following problem:

$$\begin{aligned}
\max \quad & \sum_{i=1}^{k} 0\lambda_i + \sum_{j=1}^{r} 0\theta_j \\
\text{s.t.} \quad & \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j = z \\
& \sum_{i=1}^{k} \lambda_i = 1 \\
& \lambda_i \geq 0, \qquad i = 1, \ldots, k \\
& \theta_j \geq 0, \qquad j = 1, \ldots, r
\end{aligned}$$

which is infeasible as $z \notin Q$. Its dual problem is

$$\begin{aligned} \min \quad & p^T z + q \\ \text{s.t.} \quad & p^T x^i + q \geq 0, \qquad i = 1, \ldots, k \\ & p^T w^j \geq 0, \qquad j = 1, \ldots, r \end{aligned}$$

As the latter problem has a feasible solution namely $p, q = 0$, then it is unbounded. So there exists some (p, q) such that $p^T z + q$ is negative. On the other hand, $p^T x^i + q \geq 0$ for all i and this implies that $p^T z < p^T x^i$ for all i.
Fix p as above, consider problem

$$\begin{aligned} \min \quad & p^T x \\ \text{s.t.} \quad & Ax \geq b \end{aligned}$$

If the optimal value is finite, there exists an extreme point x^i which is optimal. Since z is a feasible solution, we obtain $p^T x^i \leq p^T z$, which makes a contradiction. If the optimal cost is $-\infty$, then we must have $p^T w^j < 0$, which also makes a contradiction. ☐

Lemma 6.2. *A nonempty bounded polyhedron is the convex hull of its extreme points.*

Proof. Let

$$P = \{x \mid Ax \geq b\}$$

be a nonempty bounded polyhedron. If d is a nonzero element of the cone $C = \{x \mid Ax \geq 0\}$ and x is an element of P, we have $x + \lambda d \in P$ for all $\lambda > 0$, contradicting the boundedness of P. ☐

Lemma 6.3. *Assume that the cone $C = \{\boldsymbol{x} \mid Ax \geq 0\}$ is pointed. Then every element of C can be expressed as a nonnegative linear combination of the extreme rays of C.*