

## Lecture 14: Simplex Algorithm, Duality

Instructor: *Alex Andoni*Scribes: *Suman Mulumudi, Serena Yuan*

In these notes, we introduce duality and the Simplex algorithm (Dantzig 1949).

## 1 Simplex Algorithm

**Definition 1.**  $x$  is a vertex of the feasible region  $F \iff x$  is a basic feasible solution (bfs)  $\iff x$  tightly satisfies precisely  $n$  linearly independent constraints.

**Definition 2.** The edge  $(v, u)$  exists  $\iff v$  and  $u$  tightly satisfy  $n - 1$  shared linearly independent constraints.

**Definition 3.** The neighborhood of  $x$  is  $N(x) := \{ \text{all } y \text{ such that } (x, y) \text{ is an edge (neighbors)} \}$ .

### Simplex Algorithm:

1. Choose  $x^0 \in F$ , the starting vertex.
2. Consider  $N(x^t)$ , where  $x^t$  is  $x$  at iteration  $t$ . Choose some neighbor  $y \in N(x^t)$  such that  $c^\top y < c^\top x^t$
3. Set  $x^{t+1} = y$ , and repeat from step 2 (if such  $y$  exists). If  $y$  does not exist,  $x^t$  is optimal.

For the standard form of the LP given by

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & Ax \leq b \end{array}$$

it is possible to find a feasible starting point for Step 1 of the Simplex Algorithm by using an auxiliary LP to find  $x^0 \in F$ . This is equivalent to finding the minimum  $t$  such that  $Ax \leq b + t \cdot \mathbb{1}$ : set all  $x$ 's to be 0, and  $t$  is the most negative  $b$  value. Solve this auxiliary LP starting from  $x = 0$  and  $t = -\min_i b_i$  (which is necessarily feasible for the auxiliary LP).

- If  $t > 0$ , the original LP is not feasible
- Otherwise,  $x$  is a feasible point

Note that we obtain a point  $x$ , but if  $x$  is not a vertex of  $F$  (i.e. it does not satisfy precisely  $n$  constraints), push  $x$  to tightly satisfy precisely  $n$  constraints.

**Remark 4.** *The Simplex Algorithm will converge eventually.*

**Remark 5.** *Step 2 of the Simplex Algorithm is the "pivoting" rule.*

**Remark 6.** *Choosing  $y^* = \arg \min_{y \in N(x)} c^\top y$  is not necessarily the optimal pivoting rule (where optimal is the rule which leads to fastest convergence).*

We note that  $N(x)$  can be computed by the following algorithm.

**Pivoting Algorithm:**

- Let  $E_1, \dots, E_n$  be equalities defining  $x$
- For each  $i \in [n]$ :
  - For each  $j \in [m]$ :
    - Define  $y$  as solution to  $E_1, \dots, E_{i-1}, E_{i+1}, \dots, E_n$  and  $A_j y = b_j$
    - If  $y$  is unique and  $y \in F$ , add it to  $N(x)$ .

The total time of this algorithm is  $\# \text{ iterations} \cdot n \cdot m \cdot (GE + m \cdot n)$  where  $GE$  is the time to perform Gaussian elimination and  $m \cdot n$  is the time to check  $Ay \leq b$ .

**Remark 7.** For many pivoting rules, there exist LPs taking  $\exp(n)$  number of iterations.

**Remark 8.** "Smoothed LPs" admit polytime simple algorithms, where smoothed LPs are created by perturbing the parameters of the program. [Spielman, Teng]

**Conjecture 9.** Hirsch Conjecture: Consider graph  $G$  composed of vertices and edges (of the polytope). The graph has  $\text{poly}(n, m)$  diameter and  $\exp(n)$  vertices.

## 2 Duality

### 2.1 Reminder: Standard form of Primal LP

Previously, we defined the standard form of the primal:

**Definition 10.** The standard form of the primal LP is:

$$\begin{aligned} \text{minimize} \quad & v^* = c^\top x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

And we have shown previously that all LPs can be converted to this form.

### 2.2 The Dual

Consider the following definition of the dual:

**Definition 11.** Given a primal in standard form, the dual LP is:

$$\begin{aligned} \text{maximize} \quad & w^* = b^\top y \\ \text{subject to} \quad & A^\top y \leq c \end{aligned}$$

We will motivate this construction by demonstrating weak duality:

**Theorem 12.** Weak Duality: Given a primal LP and its optimal solution  $v^*$ , and its corresponding dual and dual optimal solution  $w^*$ , then  $v^* \geq w^*$ .

*Proof.* The proof is given in the motivation for constructing the Definition 11 above.

Consider a vector  $y \in R^m$  such that  $y^\top A \leq c^\top$  (where the inequality holds coordinate-wise). Multiplying by some  $x \in F$  where  $F$  is the feasible region of the primal yields:

$$y^\top Ax \leq c^\top x$$

Here we note we have used the fact that, in the primal, we have restricted  $x \geq 0$ . Further, noting that since  $x \in F$  it follows that  $Ax = b$ , thus can be re-written as:

$$\begin{aligned} y^\top b &= y^\top Ax \leq c^\top x \\ \implies \forall x \in F, y^\top b &\leq c^\top x \\ \implies y^\top b &\leq \min_{x \in F} c^\top x = v^* \end{aligned}$$

□

We see we have constructed the dual such that it yields an optimization problem which permits a lower bound on the optimal solution of the primal.

### 2.3 Dual of the Dual

**Theorem 13.** *For linear programs, the dual of the dual is the primal.*

*Proof.* Given the dual in Definition 11 above, the dual can be re-written in the standard form as:

$$\begin{aligned} \text{optimize} \quad w^* &= -\min_{y^+, y^-, \delta} -b^\top (y^+ - y^-) \\ \text{subject to} \quad &A_i^\top y^+ - A_i^\top y^- + \delta_i = c_i \\ &y^+, y^-, \delta \geq 0 \end{aligned}$$

Where the original  $y$  is now  $y = y^+ - y^-$ . (More details on this construction can be found in earlier notes.)

For which we can construct the dual:

$$\begin{aligned} \text{optimize} \quad v^* &= -\max_{z \in R^n} c^\top z = \min_{z \in R^n} c^\top (-z) \\ \text{subject to} \quad &Az \leq -b \\ &-Az \geq b \\ &Iz \leq 0 \end{aligned}$$

We note  $Az \leq -b$  and  $-Az \geq b$  gives  $A(-z) = b$ . Substituting  $x = -z$  recovers the primal form:

$$\begin{aligned} \text{optimize} \quad v^* &= \min_{z \in R^n} c^\top (-z) = \min_{x \in R^n} c^\top x \\ \text{subject to} \quad &Ax = b \\ &x \geq 0 \end{aligned}$$

□

## 2.4 Strong Duality

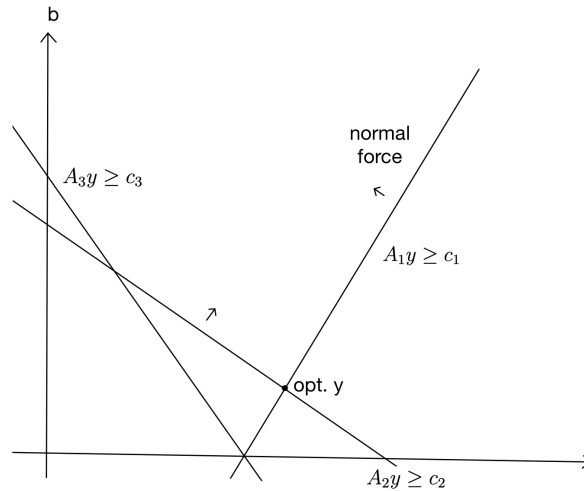
**Claim 14.** *Strong Duality: For LPs, if  $v^*$  is the solution to the primal and  $w^*$  is the solution to the dual, then  $v^* = w^*$ . That is, the primal and dual have identical solutions.*

## 2.5 Intuition for Strong Duality

Proving strong duality in general is tedious, but we will provide intuition here for why strong duality should hold by making an analogy with physics. In the dual, we are optimizing  $w^* = -\min b^\top(-y)$  subject to  $A^\top y \leq c$ . For convenience, we re-define  $-y \rightarrow y$  and  $-c \rightarrow c$ , thus the dual LP is now:

$$\begin{aligned} \text{optimize} \quad & w^* = \min b^\top y \\ \text{subject to} \quad & A^\top y \geq c \end{aligned}$$

Consider the space  $R^m$  in which  $y$  lives, and consider  $y$  to be the location of a ball. Further, orient the space such that  $b$  points upwards, and there is a gravitational force  $-b$  pulling the ball at  $y$  downwards. The constraints  $A^\top y \geq c$  create  $n$  hyperplanes at  $A_i^\top y = c$  which form a kind of "convex bowl" in the space. Thus, when the ball  $y$  is dropped into the space, it will fall until it reaches the lowest point in the bowl. Since  $b$  points upwards, this corresponds to an optimal solution to the dual problem (since the LP essentially minimizes the projection of  $y$  along  $b$ ).



If there is no plane perpendicular to  $b$ , then  $y$  will have reached a vertex of the feasible region. Returning to the physics analogy, this requires that  $y$  is supported by normal forces coming from the planes with which it is in contact. Since  $y$  is in  $R^m$ , a vertex will be found at the intersection of precisely  $m$  planes. Suppose that each hyperplane  $i$  provides a normal force of  $x_i$ . Since the ball is at rest, the total normal force must equal the gravitational force  $b$ , thus over all hyperplanes (of which  $m < n$  actually contribute):

$$\sum_i A_i x_i = b$$

Consider that  $y$  will be resting on  $m$  of  $n$  possible hyperplanes, thus  $m$  constraints will be tight. For those that are not tight, it follows that the  $x_i$  for that constraint should be 0 (since it will not be contributing to the normal force), thus  $x_i A_i y_i = 0 = c_i x_i$ . If the constraint is tight, then  $A_i y_i = c \implies x_i A_i y_i = x_i c_i$ . Noting that in the primal form, we have  $A^\top x = b$ , summing over all constraints yields:

$$\begin{aligned} \sum_i x_i A_i y &= \left( \sum_i x_i A_i \right) y = (A^\top x)^\top y = b^\top y \\ \sum_i x_i A_i y &= \sum_i c_i x_i = c^\top x \\ &\implies c^\top x = b^\top y \end{aligned}$$

Noting  $-v^* = c^\top x$  and  $-w^* = b^\top y$ , and noting that we have demonstrated weak duality  $v^* \geq w^*$ , showing an instance of  $c^\top x = b^\top y$  gives strong duality. There are cases which have not been shown (such as when a constraint creates a hyperplane perpendicular to  $b$ ), but this physical analogy should give intuition for why strong duality holds for LPs.