Lecture 10: Duality Theory

Zizhuo Wang

Institute of Data and Decision Analytics (iDDA)
Chinese University of Hong Kong, Shenzhen

Oct 10, 2018

Announcements

- Homework 3 due today
- Homework 4 due next Wednesday (10/17), at 12pm (noon)
- Office hour this week: Thursday (10/11), 10am-12pm

Review: Simplex Method

We have completed our discussions on the simplex method.

- The idea of simplex method (search among the BFS and search among the neighbors) and its justifications
- The algebraic procedures
- The simplex tableau
- Other issues about simplex method (initialization, degeneracy, etc.)

The simplex method is not a polynomial time algorithm for any configuration that people have tried.

- However, linear optimization is polynomial-time solvable.
- Later we will discuss another algorithm for LP — the interior point method, but now we turn to introduce another very important theory for linear program, the duality theory.

Duality Theory: Motivation

Consider the standard LP ($m \times n$):

$$\begin{aligned} \text{minimize}_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

We can write it as:

$$\begin{aligned} \text{minimize}_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} + \max_{\mathbf{y}\in\mathbb{R}^m} \mathbf{y}^T(\mathbf{b} - A\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \geq 0 \end{aligned}$$

Why?

- If $A\mathbf{x} \neq \mathbf{b}$, then we can find $\mathbf{y}$ such that $\mathbf{y}^T(A\mathbf{x} - \mathbf{b}) = \infty$, thus can't be optimal. So we implicitly enforce the constraint $A\mathbf{x} = \mathbf{b}$

Continue..

$$\min_{\mathbf{x}\geq 0} \quad \max_{\mathbf{y}} \quad \mathbf{c}^T\mathbf{x} + \mathbf{y}^T(\mathbf{b} - A\mathbf{x})$$

Now we assume that we can exchange the max and min (will justify it later). Then the problem becomes..

$$\max_{\mathbf{y}} \quad \mathbf{b}^T\mathbf{y} + \min_{\mathbf{x}\geq 0} \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y})$$

We claim that this is equivalent to

$$\begin{array}{ll} \text{maximize}_{\mathbf{y}} & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & A^T\mathbf{y} \leq \mathbf{c} \end{array}$$

Why?

$$\min_{\mathbf{x}\geq 0} \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = \begin{cases} 0 & \text{if } A^T\mathbf{y} \leq \mathbf{c} \\ -\infty & \text{if } A^T\mathbf{y} \not\leq \mathbf{c} \end{cases}$$

Now We Get Two Linear Programs...

$$
\begin{aligned}
\text{minimize}_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} \\
\text{subject to} \quad & A\mathbf{x} = \mathbf{b} \\
& \mathbf{x} \geq 0
\end{aligned}
$$

and

$$
\begin{aligned}
\text{maximize}_{\mathbf{y}} \quad & \mathbf{b}^T\mathbf{y} \\
\text{subject to} \quad & A^T\mathbf{y} \leq \mathbf{c}
\end{aligned}
$$

We call them dual to each other. If we call one the *primal problem*, then the other one is called *the dual problem*.

- We call $\mathbf{y}$ the dual variables (to the first LP)
- By our derivation, they should have the same optimal value (provided the exchange of min and max is valid)

Dual Problem

Now we have derived our first pair of dual problems.

- Duality theory is very important for linear optimization (and for general optimization problems as well)
- The dual problem carries much useful information for the primal problem
- It also helps to solve optimization problems

Different Duality Forms

We have derived the dual problem of the standard form. What if we want to find the dual problem of:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} \geq \mathbf{b} \end{array} \tag{1}$$

Of course, one can first transform (1) to the standard form and then derive the dual. But we can directly apply our previous arguments.

The primal problem can be written as

$$\min_{\mathbf{x}} \quad \mathbf{c}^T\mathbf{x} + \max_{\mathbf{y}\geq 0} \mathbf{y}^T(\mathbf{b} - A\mathbf{x})$$

Why?

$$\max_{\mathbf{y}\geq 0} \mathbf{y}^T(\mathbf{b} - A\mathbf{x}) = \begin{cases} 0 & \text{if } A\mathbf{x} \geq \mathbf{b} \\ \infty & \text{if } A\mathbf{x} \not\geq \mathbf{b} \end{cases}$$

Continue

Assume we can exchange the order of max and min, we get

$$\max_{\mathbf{y}\geq 0} \quad \mathbf{b}^T\mathbf{y} + \min_{\mathbf{x}} \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y})$$

This is further equivalent as

$$\begin{array}{rc} \text{maximize} & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & A^T\mathbf{y} = \mathbf{c} \\ & \mathbf{y} \geq 0 \end{array}$$

We get the dual problem of (1)

Generally

Primal			**Dual**		
minimize	$\mathbf{c}^T\mathbf{x}$		maximize	$\mathbf{b}^T\mathbf{y}$	
subject to	$\mathbf{a}_i^T\mathbf{x} \geq b_i,$	$i \in M_1,$	subject to	$y_i \geq 0,$	$i \in M_1$
	$\mathbf{a}_i^T\mathbf{x} \leq b_i,$	$i \in M_2,$		$y_i \leq 0,$	$i \in M_2$
	$\mathbf{a}_i^T\mathbf{x} = b_i,$	$i \in M_3,$		y_i free,	$i \in M_3$
	$x_j \geq 0,$	$j \in N_1,$		$A_j^T\mathbf{y} \leq c_j,$	$j \in N_1$
	$x_j \leq 0,$	$j \in N_2,$		$A_j^T\mathbf{y} \geq c_j,$	$j \in N_2$
	x_j free,	$j \in N_3,$		$A_j^T\mathbf{y} = c_j,$	$j \in N_3$

- $\mathbf{a}_i^T$ is the ith row of A, A_j is the jth column of A
- Each primal constraint corresponds to a dual variable
- Each primal variable corresponds to a dual constraint
- Equality constraints always correspond to free variables

Example

$$\begin{aligned} \text{minimize} \quad & x_1 + 2x_2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 5 \\ & x_1 - x_2 \leq 3 \\ & x_1 \geq 0, x_2 \text{ free} \end{aligned}$$

The dual problem:

- Associate constraint 1 with dual variable y_1
- Associate constraint 2 with dual variable y_2

$$\begin{aligned} \text{maximize} \quad & 5y_1 + 3y_2 \\ \text{s.t.} \quad & y_1 + y_2 \leq 1 \\ & y_1 - y_2 = 2 \\ & y_1 \geq 0, y_2 \leq 0 \end{aligned}$$

To Memorize

Primal	minimize	maximize	Dual
constraints	$\geq b_i$ $\leq b_i$ $= b_i$	≥ 0 ≤ 0 free	variables
variables	≥ 0 ≤ 0 free	$\leq c_j$ $\geq c_j$ $= c_j$	constraints

My trick to memorize:

1. Equality constraints correspond to free variables, vice versa
2. I call non-negative constraints *usual* constraints; non-positive *unusual* constraints. When maximizing, I call $\leq c_i$ *usual* constraint (upper bound for each resource), and $\geq c_i$ *unusual*; when minimizing, I call $\geq b_i$ *usual* and $\leq b_i$ *unusual*.
3. Then usual (unusual) constraints always correspond to usual (unusual) constraints.

Example

Consider the following problem:

$$\begin{array}{lrrrl}
\text{minimize} & x_1 & +2x_2 & +3x_3 & \\
\text{subject to} & -x_1 & +3x_2 & & = 5 \\
 & 2x_1 & -x_2 & +3x_3 & \geq 6 \\
 & & & x_3 & \leq 4 \\
 & x_1 \geq 0 & x_2 \leq 0 & x_3 \text{ free} &
\end{array}$$

The dual is:

$$\begin{array}{lrrrl}
\text{maximize} & 5y_1 & +6y_2 & +4y_3 & \\
\text{subject to} & -y_1 & +2y_2 & & \leq 1 \\
 & 3y_1 & -y_2 & & \geq 2 \\
 & & +3y_2 & +y_3 & = 3 \\
 & y_1 \text{ free} & y_2 \geq 0 & y_3 \leq 0 &
\end{array}$$

One More Example

Recall the support vector machine problem. The primal problem is:

$$
\begin{array}{rlr}
\text{minimize}_{\mathbf{a},b,\boldsymbol{\delta},\boldsymbol{\sigma}} & \sum_{i=1}^{n} \delta_i + \sum_{j=1}^{m} \sigma_j & \\
\text{subject to} & \mathbf{x}_i^T \mathbf{a} + b + \delta_i \geq 1, & \forall i = 1, ..., n \\
& \mathbf{y}_j^T \mathbf{a} + b - \sigma_j \leq -1, & \forall j = 1, ..., m \\
& \delta_i \geq 0, \ \sigma_j \geq 0, & \forall i = 1, ..., n, \quad j = 1, ..., m
\end{array}
$$

Associate w_i to each of the first set of constraints, v_j to each of the second set of constraints. (Suppose $\mathbf{a}$ is a d-dimensional vector.)

Example Continued

The dual problem is

$$
\begin{array}{lll}
\text{maximize}_{\mathbf{w},\mathbf{v}} & \sum_{i=1}^n w_i - \sum_{j=1}^m v_j & \\
\text{subject to} & \sum_{i=1}^n x_{ik} w_i + \sum_{j=1}^m y_{jk} v_j = 0, & \forall k = 1, ..., d \\
& \sum_{i=1}^n w_i + \sum_{j=1}^m v_j = 0 & \\
& 0 \leq w_i \leq 1, & \forall i = 1, ..., n \\
& -1 \leq v_j \leq 0, & \forall j = 1, ..., m
\end{array}
$$

If we let $X = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n]$ and $Y = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m]$, then the first constraint can also be written as

$$X\mathbf{w} + Y\mathbf{v} = \mathbf{0}$$

Example Continued

The support vector machine problem:

$$\begin{array}{rll} \text{minimize} & \sum_i \delta_i + \sum_j \sigma_j & \\ \text{subject to} & \mathbf{x}_i^T \mathbf{a} + b + \delta_i \geq 1, & \forall i \\ & \mathbf{y}_j^T \mathbf{a} + b - \sigma_j \leq -1, & \forall j \\ & \delta_i \geq 0,\ \sigma_j \geq 0, & \forall i, j \end{array}$$

Dual:

$$\begin{array}{rll} \text{maximize} & \sum_i w_i - \sum_j v_j & \\ \text{subject to} & X\mathbf{w} + Y\mathbf{v} = \mathbf{0} & \\ & \sum_i w_i + \sum_j v_j = 0 & \\ & 0 \leq w_i \leq 1, & \forall i = 1, ..., n \\ & -1 \leq v_j \leq 0, & \forall j = 1, ..., m \end{array}$$