Math 482: Linear Programming[1] Mikhail Lavrov

Lecture 30: The Primal-Dual Algorithm II

April 20, 2020 *University of Illinois at Urbana-Champaign*

1 A summary

Our goal is to solve the primal-dual pair of linear programs below:

$$(\mathbf{P}) \begin{cases} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \ge \mathbf{0} \end{cases} \qquad (\mathbf{D}) \begin{cases} \underset{\mathbf{u}\in\mathbb{R}^m}{\text{maximize}} & \mathbf{u}^\mathsf{T}\mathbf{b} \\ \text{subject to} & \mathbf{u}^\mathsf{T}A \le \mathbf{c}^\mathsf{T} \\ & \mathbf{u} \text{ unrestricted} \end{cases}$$

A summary of the primal-dual algorithm as we see it so far:

1. If necessary, tweak (**P**) and (**D**) so that $\mathbf{b} \ge \mathbf{0}$. Then, somehow start with a feasible solution $\mathbf{u}$ to the dual program (**D**).
2. Determine which constraints of (**D**) are tight, and use that to construct (**DRP**). Solve it to find $\mathbf{v}$, a direction in which to go if we want to improve $\mathbf{u}$.
3. Find the largest t such that $\mathbf{u} + t\mathbf{v}$ is still feasible for (**D**), and replace $\mathbf{u}$ by $\mathbf{u} + t\mathbf{v}$.
4. Repeat steps 2–3 until we can no longer improve $\mathbf{u}$.

Today, we will see how to improve this algorithm by solving (**RP**), the dual of (**DRP**), rather than solving (**DRP**) directly. Recall that (**RP**) and (**DRP**) look like this:

$$(\mathbf{RP}) \begin{cases} \underset{\mathbf{x}\in\mathbb{R}^n,\mathbf{y}\in\mathbb{R}^m}{\text{minimize}} & y_1 + \cdots + y_m \\ \text{subject to} & A_J\mathbf{x}_J + I\mathbf{y} = \mathbf{b} \\ & \mathbf{x},\mathbf{y} \ge \mathbf{0} \end{cases} \qquad (\mathbf{DRP}) \begin{cases} \underset{\mathbf{v}\in\mathbb{R}^m}{\text{maximize}} & \mathbf{v}^\mathsf{T}\mathbf{b} \\ \text{subject to} & \mathbf{v}^\mathsf{T}A_J \le \mathbf{c}_J^\mathsf{T} \\ & v_1, \dots, v_m \le 1 \end{cases}$$

2 Using (RP) in the primal-dual algorithm

Let's return to the example we were looking at last time:

$$(\mathbf{P}) \begin{cases} \underset{\mathbf{x}\in\mathbb{R}^3}{\text{minimize}} & 2x_1 + 2x_2 + x_3 \\ \text{subject to} & 2x_1 + x_2 - 4x_3 = 3 \\ & 4x_1 - x_2 + x_3 = 3 \\ & x_1, x_2, x_3 \ge 0 \end{cases} \qquad (\mathbf{D}) \begin{cases} \underset{\mathbf{u}\in\mathbb{R}^2}{\text{maximize}} & 3u_1 + 3u_2 \\ \text{subject to} & 2u_1 + 4u_2 \le 2 \\ & u_1 - u_2 \le 2 \\ & -4u_1 + u_2 \le 1 \end{cases}$$

[1] This document comes from the Math 482 course webpage: https://faculty.math.illinois.edu/~mlavrov/courses/482-spring-2020.html

At the point $\mathbf{u} = (1, 0)$, only the first dual constraint is tight, so (**RP**) and (**DRP**) are:

$1^T \cdot y$

$$(\mathbf{RP}) \begin{cases} \underset{\mathbf{x}\in\mathbb{R},\mathbf{y}\in\mathbb{R}^2}{\text{minimize}} & y_1 + y_2 \\ \text{subject to} & 2x_1 + y_1 = 3 \\ & 4x_1 + y_2 = 3 \\ & x_1, y_1, y_2 \geq 0 \end{cases} \qquad (\mathbf{DRP}) \begin{cases} \underset{\mathbf{v}\in\mathbb{R}^2}{\text{maximize}} & 3v_1 + 3v_2 \\ \text{subject to} & 2v_1 + 4v_2 \leq 0 \\ & v_1 \leq 1 \\ & v_2 \leq 1 \end{cases}$$

$v^T I$

2.1 Going from (RP) to (DRP)

We can solve (**RP**) in our usual way: using the simplex method. In this case, just one pivot step does the job:

	x_1	y_1	y_2	
y_1	2	1	0	3
y_2	4	0	1	3
$-z_{rp}$	−6	0	0	−6

⇝

	x_1	y_1	y_2	
y_1	0	1	$-1/2$	$3/2$
x_1	1	0	$1/4$	$3/4$
$-z_{rp}$	0	0	$3/2$	$-3/2$

$r_y^T = c_y^T - v^T A_y = 1^T - v^T$

To find the optimal solution $\mathbf{v}$ to (**DRP**), we use a modification of a method we saw earlier this semester. Recall that the reduced cost of the i^{th} variable in the tableau satisfies

$$r_i = c_i - \mathbf{v}^\mathsf{T} A_i,$$

where $\mathbf{v}$ is the dual solution. In our case, let $\mathcal{Y}$ be the positions of the variables $y_1, y_2, \ldots, y_m$. Then we have

$$\mathbf{r}_\mathcal{Y}{}^\mathsf{T} = \mathbf{c}_\mathcal{Y}{}^\mathsf{T} - \mathbf{v}^\mathsf{T} A_\mathcal{Y} = \mathbf{1}^\mathsf{T} - \mathbf{v}^\mathsf{T}$$

because (1) the costs of $y_1, \ldots, y_m$ in (**RP**) are all 1 and (2) the part of the coefficient matrix corresponding to $\mathbf{y}$ in (**RP**) is just the identity matrix.

Therefore we find the optimal solution to (**DRP**) by taking $\mathbf{v}^\mathsf{T} = \mathbf{1}^\mathsf{T} - \mathbf{r}_\mathcal{Y}{}^\mathsf{T}$. In our case, that's $\mathbf{v}^\mathsf{T} = \begin{bmatrix}1 & 1\end{bmatrix} - \begin{bmatrix}0 & \frac{3}{2}\end{bmatrix} = \begin{bmatrix}1 & -\frac{1}{2}\end{bmatrix}$. We would use this to augment $\mathbf{u}$ exactly as in the previous lecture, going from $(1, 0)$ to $(\frac{5}{3}, -\frac{1}{3})$.

$(1+t, -\frac{1}{2}t)$

At this new point $\mathbf{u} = (\frac{5}{3}, -\frac{1}{3})$, the first and second constraints of (**D**) are tight. So we get new instances of (**RP**) and (**DRP**):

~~$2u_1 + 4u_2 \leq 2$~~

$u_1 - u_2 \leq 2$ $\quad t \leq \frac{2}{3}$

$-4u_1 + u_2 \leq 1$ $\quad t \geq -\frac{10}{9}$

$\Rightarrow (\frac{5}{3}, -\frac{1}{3})$

$$(\mathbf{RP}) \begin{cases} \underset{\mathbf{x}\in\mathbb{R},\mathbf{y}\in\mathbb{R}^2}{\text{minimize}} & y_1 + y_2 \\ \text{subject to} & 2x_1 + x_2 + y_1 = 3 \\ & 4x_1 - x_2 + y_2 = 3 \\ & x_1, x_2, y_1, y_2 \geq 0 \end{cases} \qquad (\mathbf{DRP}) \begin{cases} \underset{\mathbf{v}\in\mathbb{R}^2}{\text{maximize}} & 3v_1 + 3v_2 \\ \text{subject to} & 2v_1 + 4v_2 \leq 0 \\ & v_1 - v_2 \leq 0 \\ & v_1 \leq 1 \\ & v_2 \leq 1 \end{cases}$$

2.2 Connecting instances of (RP)

We could just as easily have solved (**DRP**) directly. Why work with (**RP**)?

The reason is that when we go from $\mathbf{u} = (1, 0)$ to $\mathbf{u} = (\frac{5}{3}, -\frac{1}{3})$, the optimal $\mathbf{v}$ for the previous version of (**DRP**) is definitely not feasible for the new version of (**DRP**). After all, we kept going in our old direction $\mathbf{v}$ until we ran into an obstacle: that obstacle means that the old direction isn't usable anymore.

On the other hand, the optimal $(\mathbf{x}, \mathbf{y})$ for the previous version of (**RP**) is still feasible for the new version of (**DRP**)!

This requires some checking. First of all, it's not even clear that using the previous optimal solution even makes sense: the two instances of (**RP**) have different sets of variables. Let's look at that change more closely.

- In our example, the only thing that happens is that the new (**RP**) acquires a variable x_2 that didn't exist in the old (**RP**). That's not actually a problem: if we take the old optimal solution, and set $x_2 = 0$, we should get a feasible solution for the new (**RP**).
- However, it's also possible that some variables in the old (**RP**) disappear in the new (**RP**). This would be a problem if those variables had positive values in the old optimal solution.

However, if a variable x_i in the optimal solution to (**RP**) has a positive value, we can guarantee that it still exists in the new (**RP**). Here's why.

By complementary slackness, the constraint in (**DRP**) corresponding to x_i is tight in the optimal direction $\mathbf{v}$. For example, in the step above, $x_1 = \frac{3}{4}$ corresponds to the constraint $2v_1 + 4v_2 \le 0$, and the optimal direction $\mathbf{v} = (1, -\frac{1}{2})$ satisfies $2v_1 + 4v_2 = 0$.

That constraint goes together with a tight constraint in (**D**): in this example, with the constraint $2u_1 + 4u_2 \le 2$, which is tight at $\mathbf{u} = (1, 0)$.

When we go from $\mathbf{u}$ to $\mathbf{u} + t\mathbf{v}$, for *any* t, that constraint remains tight. In our example, $2(u_1 + tv_1) + 4(u_2 + tv_2)$ can be rewritten as $(2u_1 + 4u_2) + t(2v_1 + 4v_2)$, which simplifies to $2 + 0t = 2$. The right-hand side does not change. The same argument works generally.

3 The primal-dual algorithm with frozen variables

To make use of the observation in the previous section, let's think about solving (**RP**) differently.

We will write down a simplex tableau for (**RP**) that includes every single variable in (**P**). However, variables that are not supposed to be there are treated as "frozen". Frozen variables cannot be pivoted on and we generally ignore their existence.

Let's go back and try this again with the instance of (**RP**) we got for the dual solution $\mathbf{u} = (1, 0)$. Marking the columns of frozen variables in blue, here is what our first iteration looks like:

$$\begin{array}{c|ccccc|c} & x_1 & [\boldsymbol{x_2}] & [\boldsymbol{x_3}] & y_1 & y_2 & \\ \hline y_1 & 2 & [1] & [-4] & 1 & 0 & 3 \\ y_2 & 4 & [-1] & [1] & 0 & 1 & 3 \\ \hline -z_{rp} & -6 & [0] & [3] & 0 & 0 & -6 \end{array} \quad \rightsquigarrow \quad \begin{array}{c|ccccc|c} & x_1 & [\boldsymbol{x_2}] & [\boldsymbol{x_3}] & y_1 & y_2 & \\ \hline y_1 & 0 & [3/2] & [-9/2] & 1 & -1/2 & 3/2 \\ x_1 & 1 & [-1/4] & [1/4] & 0 & 1/4 & 3/4 \\ \hline -z_{rp} & 0 & [-3/2] & [9/2] & 0 & 3/2 & -3/2 \end{array}$$

If we ignore the frozen columns, then nothing is changing.

As we did before, we will read off the dual solution $\mathbf{v} = (1, -\frac{1}{2})$, and then go from $\mathbf{u}$ to $\mathbf{u} + \frac{2}{3}\mathbf{v} = (\frac{5}{3}, -\frac{1}{3})$. At this new $\mathbf{u}$, the first and second constraints are tight.

Now the frozen variables see some use. To solve (**RP**) in the second iteration, we simply take our final tableau from the previous iteration, and change which variables are frozen. In this case: x_1 is still unfrozen, but x_2 also becomes unfrozen.

We get the tableau on the left when we do this; pivoting on x_2, we get

	x_1	x_2	$[\boldsymbol{x_3}]$	y_1	y_2	
y_1	0	$3/2$	$[-9/2]$	1	$-1/2$	$3/2$
x_1	1	$-1/4$	$[1/4]$	0	$1/4$	$3/4$
$-z_{rp}$	0	$-3/2$	$[9/2]$	0	$3/2$	$-3/2$

$\leadsto$

	x_1	x_2	$[\boldsymbol{x_3}]$	y_1	y_2	
x_2	0	1	$[-3]$	$2/3$	$-1/3$	1
x_1	1	0	$[-1/2]$	$1/6$	$1/6$	1
$-z_{rp}$	0	0	$[0]$	1	1	0

Here, we see the typical final state of the primal-dual algorithm. The reduced costs of y_1 and y_2 are both 1; when we compute the dual solution $\mathbf{v}$, we will get $(0, 0)$. Obviously going in the direction $(0, 0)$ will not augment anything; therefore the $\mathbf{u}$ we've found is optimal.

Also, since we've gotten $y_1 + y_2 = 0$ in (**RP**), the primal solution with x_1 and $x_2 = 1$ we see in this tableau is the optimal solution to (**P**). It satisfies $A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$, and it satisfies complementary slackness: whenever a dual constraint is slack (whenever $\mathbf{u}^\mathsf{T} A_j < c_j$) we have $x_j = 0$.