Math 3272: Linear Programming[1]	Mikhail Lavrov
Lecture 15: Duality and the simplex method	
October 6, 2022	*Kennesaw State University*

1 Finding the dual solution from the dictionary

In the previous lecture, we discussed complementary slackness, which lets us find the optimal dual solution given the optimal primal solution. This still involves a little bit of work solving the equations.

In practice, we might have an optimal dictionary and not just an optimal primal solution. If this is the case, then there is a more concrete expression we can give for the dual solution. Actually, there are two ways to find the dual solution:

- one general way that always works;
- one very quick method that works for linear programs that started in $A\mathbf{x} \le \mathbf{b}$ form and added slack variables.

1.1 A general formula

The simplex method is applied to problems in equational form. Our dual in this case looks like the following, with $\mathbf{u}$ unconstrained (each u_i can be positive or negative):

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

Recall that we have a formula for the dictionaries we get through the simplex method. If we choose basic variables $\mathcal{B}$ and nonbasic variables $\mathcal{N}$, then the corresponding dictionary is

$$\begin{array}{c} \zeta \;=\; \mathbf{u}^\mathsf{T}\mathbf{b} + \Big((\mathbf{c}_\mathcal{N})^\mathsf{T} - \mathbf{u}^\mathsf{T}A_\mathcal{N}\Big)\mathbf{x}_\mathcal{N} \\ \hline \mathbf{x}_\mathcal{B} = (A_\mathcal{B})^{-1}\mathbf{b} - (A_\mathcal{B})^{-1}A_\mathcal{N}\mathbf{x}_\mathcal{N} \end{array}$$

where $\mathbf{u}^\mathsf{T} = \mathbf{c}_\mathcal{B}{}^\mathsf{T}(A_\mathcal{B})^{-1}$. It is not a coincidence that the vector $\mathbf{u}$ used in this formula was given the same letter as the vector $\mathbf{u}$ we are using for the dual solution: they are the same!

More precisely, suppose that we have achieved an *optimal* dictionary for maximizing $\mathbf{c}^\mathsf{T}\mathbf{x}$. This means that our reduced costs are all less than or equal to 0: we have no variables left worth pivoting on. In other words, $(\mathbf{c}_\mathcal{N})^\mathsf{T} - \mathbf{u}^\mathsf{T}A_\mathcal{N} \le 0$, or $\mathbf{u}^\mathsf{T}A_\mathcal{N} \ge (\mathbf{c}_\mathcal{N})^\mathsf{T}$. This looks a lot like the constraints in **(D)**: more precisely, it *is* the constraints, but only the ones indexed by $\mathcal{N}$.

What about the constraints indexed by $\mathcal{B}$? These constraints correspond to the basic variables, which are probably positive in our optimal solution, so we expect them to be satisfied with equality: we expect that $\mathbf{u}^\mathsf{T}A_\mathcal{B} = (\mathbf{c}_\mathcal{B})^\mathsf{T}$. This is also true, since $\mathbf{u}^\mathsf{T} = (\mathbf{c}_\mathcal{B})^\mathsf{T}(A_\mathcal{B})^{-1}$.

[1]This document comes from an archive of the Math 3272 course webpage: `http://misha.fish/archive/3272-fall-2022`

1.2 Strong duality

The formula we've just found has a theoretic use and not just a practical one. We can use it to show that when **(P)** has an optimal solution $\mathbf{x}$, the dual solution $\mathbf{u}$ we find satisfies $\mathbf{u}^\mathsf{T}\mathbf{b} = \mathbf{c}^\mathsf{T}\mathbf{x}$: the primal and dual solutions have the same objective value. To prove this, we need to remember our formula for the primal optimal solution we read off from the dictionary: we set $\mathbf{x}_\mathcal{N} = \mathbf{0}$, and get $\mathbf{x}_\mathcal{B} = (A_\mathcal{B})^{-1}\mathbf{b}$. Therefore

$$\mathbf{c}^\mathsf{T}\mathbf{x} = (\mathbf{c}_\mathcal{B})^\mathsf{T}\mathbf{x}_\mathcal{B} + (\mathbf{c}_\mathcal{N})^\mathsf{T}\mathbf{x}_\mathcal{N} = (\mathbf{c}_\mathcal{B})^\mathsf{T}(A_\mathcal{B})^{-1}\mathbf{b} + (\mathbf{c}_\mathcal{N})^\mathsf{T}\mathbf{0} = \mathbf{u}^\mathsf{T}\mathbf{b}.$$

This is a proof of *strong* duality in a special case:

Theorem 1 (Strong duality)**.** *Whenever* ***(P)*** *has an optimal solution* $\mathbf{x}$*, it is also true that* ***(D)*** *has an optimal solution* $\mathbf{u}^\mathsf{T}$ *with the same objective value.*

The general case of strong duality can be deduced from this one, since all linear programs can be put into equational form. (The proof is not automatic, since when we have a linear program in two forms, its dual also has two forms, so optimal dual solutions also look different. We would need to check that the dual solution we got from the simplex method can be used to "recover" a dual solution for the dual of the original linear program.)

1.3 A special case

It is worthwhile to look at one other case for **(P)**: where we start with inequalities $A\mathbf{x} \le \mathbf{b}$, and add slack variables to put it in equational form. Written as a matrix equation, the equational form of $A\mathbf{x} \le \mathbf{b}$ is $A\mathbf{x} + I\mathbf{w} = \mathbf{b}$. Something silly happens with the dual when we make this change:

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

When we add slack variables, **(D)** still has nonnegativity constraints on $\mathbf{u}$, but instead of being treated separately as nonnegativity constraints, they are simply the constraints corresponding to the primal variables $\mathbf{w}$.

Looking at our dictionary formula, we can notice that if x_i is a nonbasic variable, then its reduced cost x_i is given by $c_i - \mathbf{u}^\mathsf{T}A_i$: the right-hand side of the dual constraint corresponding to x_i, minus the left-hand side of that constraint. This is also true if x_i is a basic variable, assuming that we consider the reduced cost of a basic variable to be 0.

What if we do this for a slack variable? The dual constraint corresponding to slack variable w_i is just the constraint $u_i \ge 0$. The right-hand side minus the left-hand side is just equal to $-u_i$. So we deduce a simplified rule for finding $\mathbf{u}^\mathsf{T}$:

Theorem 2. *If* ***(P)*** *started out in the inequality form* $A\mathbf{x} \le \mathbf{b}$*, then an optimal solution* $\mathbf{u}$ *for* ***(D)*** *can be read off from the optimal dictionary for* ***(P)*** *by taking the negatives of the reduced costs of the slack variables.*

1.4 Examples

Let's start with an example in equational form. Take the following primal-dual pair:

$$
\textbf{(P)}\begin{cases}
\underset{\mathbf{x}\in\mathbb{R}^4}{\text{maximize}} & x_1 \phantom{{}+x_2} + 2x_3 - x_4 \\
\text{subject to} & x_1 + x_2 + x_3 + x_4 = 4 \quad (u_1) \\
& x_1 + 2x_2 + 3x_3 + 4x_4 = 10 \quad (u_2) \\
& x_1, x_2, x_3, x_4 \ \geq\ 0
\end{cases}
\qquad
\textbf{(D)}\begin{cases}
\underset{u_1,u_2\in\mathbb{R}}{\text{minimize}} & 4u_1 + 10u_2 \\
\text{subject to} & u_1 + u_2 \geq 1 \quad (x_1) \\
& u_1 + 2u_2 \geq 0 \quad (x_2) \\
& u_1 + 3u_2 \geq 2 \quad (x_3) \\
& u_1 + 4u_2 \geq -1 \quad (x_4)
\end{cases}
$$

To get started finding the optimal dual solution, it's enough for me to tell you that in the optimal primal solution, x_1 and x_3 are basic; you don't even need to know what their values are! Then we use the formula $\mathbf{u}^\mathsf{T} = (\mathbf{c}_\mathcal{B})^\mathsf{T}(A_\mathcal{B})^{-1}$ to compute

$$
\begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}.
$$

Therefore $(u_1, u_2) = (\frac{1}{2}, \frac{1}{2})$ is the optimal solution.

Now let's look at an example with slack variables. Here, we'll actually need to know the optimal dictionary; on the other hand, we will not have to do matrix inverse calculations. Take the following example:

$$
\textbf{(P)}\begin{cases}
\underset{x,y\in\mathbb{R}}{\text{maximize}} & 2x + 3y \\
\text{subject to} & -x + y \leq 3 \quad (u_1) \\
& x - 2y \leq 2 \quad (u_2) \\
& x + y \leq 7 \quad (u_3) \\
& x, y \ \geq 0
\end{cases}
\qquad
\textbf{(D)}\begin{cases}
\underset{u_1,u_2,u_3\in\mathbb{R}}{\text{minimize}} & 3u_1 + 2u_2 + 7u_3 \\
\text{subject to} & -u_1 + u_2 + u_3 \geq 2 \quad (x) \\
& u_1 - 2u_2 + u_3 \geq 3 \quad (y) \\
& u_1, u_2, u_3 \ \geq 0
\end{cases}
$$

We add slack variables to **(P)**, solve it, and end up at the following optimal dictionary:

$$
\begin{array}{rl}
\max \zeta = & 19 - \frac{1}{2}w_1 - \frac{5}{2}w_3 \\
\hline
x = & 2 + \frac{1}{2}w_1 - \frac{1}{2}w_3 \\
y = & 5 - \frac{1}{2}w_1 - \frac{1}{2}w_3 \\
w_2 = & 10 - \frac{3}{2}w_1 - \frac{1}{2}w_3
\end{array}
$$

The reduced costs of w_1 and w_3 are $-\frac{1}{2}$ and $-\frac{5}{2}$, telling us that in the dual optimal solution, $u_1 = \frac{1}{2}$ and $u_3 = \frac{5}{2}$. What about w_2? It's a basic variable, so its reduced cost is automatically 0. Therefore $(u_1, u_2, u_3) = (\frac{1}{2}, 0, \frac{5}{2})$ is an optimal solution to **(D)**.

2 The dual simplex method

It would be fair to complain: in all of these examples, once we've solved **(P)**, why do we care about the values of **(D)**? We will see some meaning to those values in future lectures. Today, we will look at a surprising use that the dual solution has, even when we *don't* care what it is.

Here (on the left) is a linear program we looked at earlier in the semester:

$$\begin{array}{rl} \underset{x_1,x_2\in\mathbb{R}}{\text{minimize}} & 4.5x_1 + 3x_2 \\ \text{subject to} & x_1 + x_2 \ge 5 \\ & 3x_1 + x_2 \ge 7 \\ & x_1 + 2x_2 \ge 6 \\ & x_1, x_2 \ \ge 0 \end{array} \qquad \begin{array}{rcl} \min \zeta = & 0 + 4.5x_1 + 3x_2 \\ \hline w_1 = & -5 + x_1 + x_2 \\ w_2 = & -7 + 3x_1 + x_2 \\ w_3 = & -6 + x_1 + 2x_2 \end{array}$$

On the right is a *very bad* initial dictionary for it. It is not feasible: every single basic variable has a negative value!

But let's look at the bright side: all the reduced costs are positive! This is just what we want to see in a minimization problem. It would indicate that we've found an optimal solution... if it weren't for that pesky "not actually feasible" problem...

Now that we know about finding dual solutions from the dictionary, we know that these reduced costs are exactly the information we need to know that the dual solution we can extract from it is feasible. More precisely, the dual solution here has $(u_1, u_2, u_3) = (0, 0, 0)$; the dual constraints are $u_1 + 3u_2 + u_3 \le 4.5$ and $u_1 + u_2 + 2u_3 \le 3$, and they are satisfied with a slack of 4.5 and 3, which are precisely the reduced costs in our dictionary. Our "optimal-but-not-feasible" solution to the primal corresponds to a feasible (but not optimal) dual solution!

The **dual simplex method** takes this idea and runs with it. Call a dictionary **dual feasible** if all the reduced costs are the correct sign for optimality. We will start with a dual feasible dictionary, and do pivot steps that preserve dual feasibility, while getting the dictionary closer to ordinary (primal) feasibility. To do this, the overall strategy is: **choose a basic variable whose value is negative to leave the basis, then choose an entering variable so that dual feasibility is preserved.**

In this example, we're spoiled for choice in leaving variables: all three of w_1, w_2, w_3 are negative. Let's pick w_1 for no good reason. Meanwhile, we don't know how to choose an entering variable yet, so let's try both. Here are the two dictionaries we can get if either x_1 (left) or x_2 (right) enters the basis:

$$\begin{array}{rl} \min \zeta = & 22.5 + 4.5w_1 - 1.5x_2 \\ \hline x_1 = & 5 + w_1 - x_2 \\ w_2 = & 8 + 3w_1 - 2x_2 \\ w_3 = & -1 + w_1 + x_2 \end{array} \qquad \begin{array}{rl} \min \zeta = & 15 + 1.5x_1 + 3w_1 \\ \hline x_2 = & 5 - x_1 + w_1 \\ w_2 = & -2 + 2x_1 + w_1 \\ w_3 = & 4 - x_1 + 2w_1 \end{array}$$

Choosing x_1 is bad: we end up losing dual feasibility. On the other hand, dual feasibility is preserved if we pivot on x_2. What are the rules we have to follow to make this decision in general?

1. First, we have to pick an entering variable with **a positive coefficient in the leaving variable's equation**. In this example, both x_1 and x_2 had this property, so we didn't notice.

   The reason for this rule is to make sure that our leaving variable ends up with the correct sign of reduced cost. When we did the substitution of either x_1 or x_2 in ζ's equation, the old reduced cost was multiplied by the coefficient of w_1, so that coefficient had to be positive.

   This is our "shortlist for entering variables", analogous to the "shortlist for leaving variables" in the ordinary simplex method.

2. When multiple entering variables satisfy this property, we should compare ratios. Specifically, we compute the ratio

$$\frac{\text{reduced cost of variable}}{\text{coefficient in leaving variable's row}}$$

and choose the entering variable with the smallest ratio. This is the calculation we get if we track what happens to the reduced costs of other variables, and make sure that they stay positive.

In a maximization problem, dual feasibility means negative reduced costs, and we want to keep them negative. In that case, all these ratios will be negative, and we want to pick the least negative ratio (the one closest to 0). In other words, if we take absolute values first, the rule stays the same.

Let's do another pivot from the dictionary where x_2 is a basic variable. The only negative basic variable in that dictionary is w_2, so let's make w_2 the leaving variable to fix that. In the equation $w_2 = -2 + 2x_1 + w_1$, both x_1 and w_1 have positive coefficients. We compute the ratios: x_1's ratio is $\frac{1.5}{2} = 0.75$ and w_1's ratio is $\frac{3}{1} = 3$. This means x_1 should be our entering variable, since its ratio is smaller.

Our new dictionary is

$$\begin{array}{rrrr}
\multicolumn{4}{c}{\min \zeta = 16.5 + 0.75w_2 + 2.25w_1} \\
\hline
x_2 = & 4 - & 0.5w_2 + & 1.5w_1 \\
x_1 = & 1 + & 0.5w_2 - & 0.5w_1 \\
w_3 = & 3 - & 0.5w_2 + & 2.5w_1
\end{array}$$

and it is both feasible and dual feasible. So we've found the optimal solution! It is $(x_1, x_2) = (1, 4)$ with objective value 16.5.

We will see many uses of the dual simplex method in the future, but there is one practical use we can see directly from this example. If we used the ordinary simplex method, we would have had to do two phases, because we don't have an initial basic feasible solution! Meanwhile, we *do* have a initial basic solution which is dual feasible, so the dual simplex method is easier to start.

Under the hood, the dual simplex method is actually applying the simplex method to the dual linear program. However, we don't have to know that to know what is going on: we don't have to know what the dual constraints are, or the values of the dual variables.