Lecture 14: Complementary slackness

October 4, 2022 *Kennesaw State University*

1 A solution we suspect to be optimal

1.1 A shipping problem

Problem 1. *You own two chocolate stores: one in Atlanta, and one in Seattle. They buy chocolate chips from three different factories. The store in Atlanta is bigger, and buys* 10 *pounds of chocolate chips a day; the store in Seattle only buys* 5 *pounds of chocolate chips a day.*

Since the two stores are very far apart, shipping costs are very different. They are given by the table below:

	from Factory #1	*from Factory #2*	*from Factory #3*
to Atlanta	*\$7/lb*	*\$12/lb*	*\$10/lb*
to Seattle	*\$10/lb*	*\$12/lb*	*\$20/lb*

Additionally, each factory can ship at most 6 *pounds of chocolate chips per day (total).*

What is the most cost-efficient way to supply both stores with chocolate chips?

To model this linear program, our first step is to understand the variables. What quantities do we need to know to specify how we're supplying both stores? We need a variable telling us how many pounds of chocolate chips are shipped from *each* factory to *each* store.

Let's write a_1, a_2, a_3 for the amount shipped from factories 1, 2, 3 respectively to Atlanta, and s_1, s_2, s_3 for the amount shipped from factories 1, 2, 3 respectively to Seattle. These are all non-negative variables.

We have two "demand constraints": each store needs a certain amount of chocolate. We can write these as $a_1+a_2+a_3 = 10$ and $s_1+s_2+s_3 = 5$. We also have three "supply constraints": each factory can ship at most 6 pounds of chocolate per day. We can write these as $a_1 + s_1 \le 6$, $a_2 + s_2 \le 6$, and $a_3 + s_3 \le 6$. We must minimize the total cost of shipping, which we can get by multiplying the cost per pound in each entry of the table by the amount shipped from that factor to that store.

This gives us the primal linear program **(P)** below:

$$
\textbf{(P)} \begin{cases}
\underset{\mathbf{a},\mathbf{s}\in\mathbb{R}^3}{\text{minimize}} & 7a_1 + 12a_2 + 10a_3 + 10s_1 + 12s_2 + 20s_3 \\
\text{subject to} & a_1 + a_2 + a_3 = 10 \quad (u_1) \\
& s_1 + s_2 + s_3 = 5 \quad (u_2) \\
& a_1 + s_1 \le 6 \quad (v_1) \\
& a_2 + s_2 \le 6 \quad (v_2) \\
& a_3 + s_3 \le 6 \quad (v_3) \\
& a_1, a_2, a_3, s_1, s_2, s_3 \ge 0
\end{cases}
$$

[1]This document comes from the Math 3272 course webpage: https://facultyweb.kennesaw.edu/mlavrov/courses/3272-fall-2022.php

1.2 Taking the dual

To practice what we learned in the previous lecture, and to prepare for the next question, let's find the dual of this linear program. I've chosen variables u_1, u_2 for the two supply constraints and v_1, v_2, v_3 for the three demand constraints; these are written above in parentheses next to the constraint they "own".

Because the primal is a minimization problem, the dual will be a maximization problem: we will maximize the lower bound on the cost in **(P)** that we can prove by combining its constraints. Mechanically applying the rules in the previous lecture, we can derive the following dual:

$$
\textbf{(D)} \begin{cases} \underset{\mathbf{u}\in\mathbb{R}^2, \mathbf{v}\in\mathbb{R}^3}{\text{maximize}} & 10u_1 + 5u_2 + 6v_1 + 6v_2 + 6v_3 \\ \text{subject to} & u_1 + v_1 \le 7 \quad (a_1) \\ & u_1 + v_2 \le 12 \quad (a_2) \\ & u_1 + v_3 \le 10 \quad (a_3) \\ & u_2 + v_1 \le 10 \quad (s_1) \\ & u_2 + v_2 \le 12 \quad (s_2) \\ & u_2 + v_3 \le 20 \quad (s_3) \\ & v_1, v_2, v_3 \le 0 \end{cases}
$$

Let's also try to understand how these lower bounds work, so that we can better understand those rules.

A working lower bound for **(P)** would be an inequality $Pa_1 + Qa_2 + Ra_3 + Ss_1 + Ts_2 + Us_3 \ge X$, where P, Q, R, S, T, U are less than the costs $7, 12, 10, 10, 12, 20$ respectively. This would make $Pa_1 + Qa_2 + Ra_3 + Ss_1 + Ts_2 + Us_3$ a lower bound on the primal objective function $7a_1 + 12a_2 + 10a_3 + 10s_1 + 12s_2 + 20s_3$, which means X would also be a lower bound on that primal objective function. (Since all variables in **(P)** are nonnegative, it's okay if some coefficients P, Q, R, S, T, U are less than their corresponding costs.) This gives us all six constraints in **(D)**.

The objective function in **(D)** comes from seeing what the lower bound X will be if we multiply the constraints in **(P)** by coefficients u_1, u_2, v_1, v_2, v_3 and add them up. Since we want the most informative (and therefore greatest) lower bound, we want to maximize.

The trickiest part is understanding the types of variables **(D)** has. I've written $v_1, v_2, v_3 \le 0$, and this is not a typo: v_1, v_2, v_3 really are *nonpositive* variables. Why? It's because the corresponding constraints $a_1 + s_1 \le 6$, $a_2 \le s_2 \le 6$, and $a_3 + s_3 \le 6$ are all "$\le$" inequalities: they give upper bounds. To turn them into lower bounds we want, we need to multiply them by a negative number to flip them.

Similarly, u_1, u_2 are unconstrained, because the equations can be multiplied by any coefficient: positive or negative.

Here's a few feasible solutions to **(D)** to look at. First, suppose we take $u_1 = 7$, $u_2 = 10$, and $v_1 = v_2 = v_3 = 0$. This correspond to adding together $7a_1 + 7a_2 + 7a_3 = 70$ and $10s_1 + 10s_2 + 10s_3 = 50$ to get

$$7a_1 + 7a_2 + 7a_3 + 10s_1 + 10s_2 + 10s_3 = 120.$$

Since the objective function $7a_1 + 12a_2 + 10a_3 + 10s_1 + 12s_2 + 20s_3$ is at least as big as the left-hand-side of the equation above, 120 is a lower bound on the objective value.

If we increased u_1 to 10, then that wouldn't be true: the coefficient of a_1 would be too big. But supposed we fixed that by setting $v_1 = -3$: subtracting $3a_1 + 3s_1 \le 18$. We'd get:

$$10(a_1 + a_2 + a_3) + 10(s_1 + s_2 + s_3) - 3(a_1 + s_1) \ge 10 \cdot 10 + 10 \cdot 5 - 3 \cdot 6$$
$$7a_1 + 10a_2 + 10a_3 + 7s_1 + 10s_2 + 10s_3 \ge 132.$$

This lets us deduce a better lower bound!

1.3 An example of complementary slackness

Problem 2. *Adding on to the previous problem, suppose that historically, the store in Atlanta opened first. You did the clear optimal thing and bought* 6 *pounds of chocolate chips from factory #1 (with the cheapest price) and* 4 *pounds from factory #3 (with the second-cheapest price).*

Then, the store in Seattle opened. Since factory #1 has no more chocolate chips, you decided to ship 5 *pounds from factory #2. This seems reasonable, but now you're not sure. Is this the most cost-effective way to supply both stores?*

In other words, is the solution $(a_1, a_2, a_3, s_1, s_2, s_3) = (6, 0, 4, 0, 5, 0)$, with objective value 142, optimal?

We already have some ways of checking that. We could try to find a dictionary which has this as its basic feasible solution, for example, and find the reduced costs. But let's explore another option. If there is a feasible solution to **(D)** with objective value 142, then that would prove that we've found the optimal solution to **(P)**.

In fact, by looking at our solution to **(P)**, we can make some deductions about what **(D)** has to do. They come in two types.

Deduction 1. Suppose that our dual solution manages to prove the inequality

$$Pa_1 + Qa_2 + Ra_3 + Ss_1 + Ts_2 + Us_3 \ge 142$$

for some P, Q, R, S, T, U. In general, this inequality only needs to have $P \le 7$, $Q \le 12$, and so on, to be a lower bound on **(P)**'s objective function. However, we can argue that actually, since $a_1 = 6$ in the primal solution, its coefficient P must be *exactly* 7. If the coefficient of P were a smaller number like 6.5, it would mean that our dual solution would still prove a lower bound of 142 when the price of shipping from factory #1 to Atlanta dropped to \$6.50 per pound. But that's impossible, since we know that our solution to **(P)** gets cheapper by \$3 in that case!

Similarly, the coefficients of a_3 and s_2 must be exact. This tells us that three of the inequalities in **(D)** must actually be equations if we are to match the bound of 142: we must get $u_1 + v_1 = 7$, $u_1 + v_3 = 10$, and $u_2 + v_2 = 12$.

Deduction 2. In our solution to **(P)**, the constraint $a_2 + s_2 \le 6$ is **slack**: actually, $a_2 + s_2 = 0 + 5 < 6$. This means that if we *use* this constraint to prove an inequality

$$Pa_1 + Qa_2 + Ra_3 + Ss_1 + Ts_2 + Us_3 \ge 142,$$

then for our solution to **(P)**, it will actually prove a strict inequality with $<$. This is impossible: it would prove that our primal solution has objective value strictly less than 142, which is false.

Therefore we shouldn't use that constraint in our hypothetical lower bound of 142: we should have $v_2 = 0$ in the dual solution we want. Similarly, since $a_3 + s_3 \le 6$ is slack in our solution to **(P)**, we should have $v_3 = 0$ in the solution to **(D)** we're looking for.

Combining the two deductions: since $u_2 + v_2 = 12$ and $v_2 = 0$, we want $u_2 = 12$. Since $u_1 + v_3 = 10$ and $v_3 = 0$, we want $u_1 = 10$. Finally, since $u_1 + v_1 = 7$ and $u_1 = 10$, we want $v_1 = -3$.

This is all resting on the hypothetical assumption that our solution to **(P)** is optimal and has a matching lower bound based on a solution to **(D)**. So it is **extremely important** to check our work: is the resulting dual solution $(u_1, u_2, v_1, v_2, v_3) = (10, 12, -3, 0, 0)$ actually a feasible solution for **(D)**?

It turns out that yes: this solution satisfies all six constraints in **(D)**, and has an objective value of $10 \cdot 10 + 12 \cdot 5 - 3 \cdot 6 = 142$. Therefore our primal solution is optimal: the shipping plan does not need to be changed!

(If we had started with a suboptimal solution to **(P)**, we would have gotten a dual solution that fails this final check; that's why checking is so important.)

2 Complementary slackness

The technique we used in the problem above is called **complementary slackness**. Complementary slackness is a limitation on what can happen if we have a feasible solution to **(P)** and a feasible solution to **(D)** with the same objective value (in which case they're both optimal). In words, it says the following:

- Whenever our feasible solution to **(P)** has a slack constraint (the two sides of the inequality are not equal), the corresponding dual variable must be 0 in our feasible solution to **(D)**.

  In other words, whenever a dual variable is not zero, the corresponding primal constraint must be tight: the two sides must be equal.

- Whenever our feasible solution to **(D)** has a slack constraint (the two sides of the inequality are not equal), the corresponding primal variable must be 0 in our feasible solution to **(P)**.

  In other words, whenever a primal variable is not zero, the corresponding dual constraint must be tight: the two sides must be equal.

The proof is just the sort of reasoning we used in our deductions above, but generalized. Let's consider one specific case: when the primal and dual have the form

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

The proof is not significantly different in all other cases, there are just a lot of cases to check.

Theorem 1 (Complementary slackness)**.** *Suppose that we have a feasible solution* $\mathbf{x}$ *for* ***(P)*** *and a feasible solution* $\mathbf{u}^\mathsf{T}$ *for* ***(D)*** *with* $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{u}^\mathsf{T}\mathbf{b}$*. Then the following relationship holds:*

- *For all i, either $(A\mathbf{x})_i = b_i$ or $u_i = 0$.*
- *For all j, either $x_j = 0$ or $(\mathbf{u}^\mathsf{T} A)_j = c_i$.*

Proof. Recall our proof of weak duality: we showed that $\mathbf{c}^\mathsf{T}\mathbf{x} \le \mathbf{u}^\mathsf{T}\mathbf{b}$ by showing that $\mathbf{c}^\mathsf{T}\mathbf{x} \le \mathbf{u}^\mathsf{T}A\mathbf{x} \le \mathbf{u}^\mathsf{T}\mathbf{b}$. So if $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{u}^\mathsf{T}\mathbf{b}$, then we must have equality in the second equation as well: $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{u}^\mathsf{T}A\mathbf{x} = \mathbf{u}^\mathsf{T}\mathbf{b}$.

We can rewrite $\mathbf{u}^\mathsf{T}A\mathbf{x} = \mathbf{u}^\mathsf{T}\mathbf{b}$ as $\mathbf{u}^\mathsf{T}(\mathbf{b} - A\mathbf{x}) = 0$. This is a dot product which we can expand as a sum: we must have

$$\sum_{i=1}^{m} u_i(b_i - (A\mathbf{x})_i) = 0.$$

In every term, we must have $u_i \ge 0$ (since $\mathbf{u}$ is feasible for **(D)**) and $b_i - (A\mathbf{x})_i \ge 0$ (since $\mathbf{x}$ is feasible for **(P)**). So every term of the sum is nonnegative, and the only way for the sum to be 0 is to have every term equal to 0. Therefore for all i, $u_i(b_i - (A\mathbf{x})_i) = 0$, which means that either $(A\mathbf{x})_i = b_i$ or $u_i = 0$.

This proves the first bullet point. For the second bullet point, we use the same reasoning, but applied to the equation $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{u}^\mathsf{T}A\mathbf{x}$, rewritten as $(\mathbf{u}^\mathsf{T}A - \mathbf{c}^\mathsf{T})\mathbf{x} = 0$. □

As we saw in today's example, complementary slackness can be useful when we have a candidate solution, and we want to know whether it is optimal. (Note that if we find a feasible solution $\mathbf{x}$ to **(P)** and a feasible solution $\mathbf{u}$ to **(D)** such that $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{u}^\mathsf{T}\mathbf{b}$, then weak duality automatically tells us that both solutions are optimal!)