4 LP Duality

One of the most important ideas in linear programming is duality. The LP duality can be extended to all convex optimization problems, in which case it is often referred to as **Lagrangian duality**. The idea of duality is based on a very basic algorithmic idea: An optimization can be solved by introducing and updating an upper bound f_u and a lower bound f_l for the optimal cost f^* i.e., $f_l \leq f^* \leq f_u$. If at any step $f_l = f_u$, then the optimality is obtained. Very often in practice, the lower and the upper bounds cannot be exactly equal to each other for numerical reasons. Hence, an algorithm may stop when the gap between them is small enough. Consider a generic minimization problem. Finding an upper bound for this problem is simple: just take any feasible solution $\mathbf{x} \in \Omega$. Then, from the definition of the optimal value, we have that $f^* \leq f(\mathbf{x})$. Hence $f_u = f(\mathbf{x})$ is an upper bound. By this definition, if one needs to improve the upper bound, one can simply find another feasible solution with a smaller cost. Finding and updating lower bounds is more complicated and is what we call dual. There is a general idea in finding lower bounds for LPs, which we explain in the following example:

Example 10. Consider the following optimization:

Here is the same steps in a more abstract form. Let us say that an optimization is written as

$$
\begin{gathered}
f^* = \min_{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n} c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \\
\text{subject to} \\
a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \leq b_1 \\
a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n \leq b_2 \\
\vdots \\
a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \leq b_m
\end{gathered} \tag{39}
$$

Introduce nonpositive **dual variables** $u_1, u_2, \ldots, u_m \leq 0$. Multiply the k^{th} inequality constraint in (39) to u_k and add the resulting inequalities, which gives

$$
\left(\sum_{k=1}^{m} u_k a_{k1}\right) x_1 + \left(\sum_{k=1}^{m} u_k a_{k2}\right) x_2 + \ldots + \left(\sum_{k=1}^{m} u_k a_{kn}\right) x_n \geq \sum_{k=1}^{n} u_k b_k \tag{40}
$$

If the dual variables are chosen, such that

$$
\left(\sum_{k=1}^{m} u_k a_{k1}\right) = c_1, \quad \left(\sum_{k=1}^{m} u_k a_{k2}\right) = c_2, \quad \ldots \quad , \left(\sum_{k=1}^{m} u_k a_{kn}\right) = c_n \tag{41}
$$

Then, we conclude that for all feasible solution $\mathbf{x}$, the relation $f(\mathbf{x}) \geq \sum_{k=1}^{n} u_k b_k$. Hence the optimal value also satisfies $f^* \geq \sum_{k=1}^{n} u_k b_k$, which using the matrix notation can be written as $f^* \geq \mathbf{u}^T \mathbf{b}$. We can also use the matrix notation and write (41) as

$$
\mathbf{A}^T \mathbf{u} = \mathbf{c} \tag{42}
$$

We can summarize what we have found in the following statement: *Denote the optimal value in* (30) *by* f^*. *If a nonpositive vector* $\mathbf{u} \leq 0$ *satisfies* $\mathbf{A}^T \mathbf{u} = \mathbf{c}$, *then* $f^* \geq \mathbf{u}^T \mathbf{b}$.

Now, I can take a step further: If any vector $\mathbf{u} \leq 0$ satisfying $\mathbf{A}^T \mathbf{u} = \mathbf{c}$ provides me a lower bound $\mathbf{u}^T \mathbf{b}$, then I can take the maximum of such lower bounds as the "best" lower bound. The

$\min_{\mathbf{x}} \mathbf{c}^T\mathbf{x}$ subject to $\mathbf{A}\mathbf{x} \leq \mathbf{b}$	$\max_{\mathbf{u}} \mathbf{b}^T\mathbf{u}$ subject to $\mathbf{A}^T\mathbf{u} = \mathbf{c}$ $\mathbf{u} \leq \mathbf{0}$
$\min_{\mathbf{x}} \mathbf{c}^T\mathbf{x}$ subject to $\mathbf{D}\mathbf{x} = \mathbf{e}$ $\mathbf{x} \geq 0$	$\max_{\mathbf{u}} \mathbf{e}^T\mathbf{u}$ subject to $\mathbf{D}^T\mathbf{u} \geq \mathbf{c}$
$\min_{\mathbf{x}} \mathbf{c}^T\mathbf{x}$ subject to $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ $\mathbf{x} \geq 0$	$\max_{\mathbf{u}} \mathbf{b}^T\mathbf{u}$ subject to $\mathbf{A}^T\mathbf{u} \leq \mathbf{c}$ $\mathbf{u} \leq \mathbf{0}$

Table 1: The primal and dual relations.

maximization leads to the following optimization:

$$\begin{aligned} &\max_{\mathbf{u}\in\mathbb{R}^m} \mathbf{u}^T\mathbf{b} \\ &\text{subject to} \\ &\mathbf{A}^T\mathbf{u} = \mathbf{c} \\ &\mathbf{u} \leq 0 \end{aligned} \tag{43}$$

The optimization is (43) is called the dual of the optimization in (30). In this context, the optimization in (30) is referred to as **primal**. We derived the dual optimization for the specific form in (30). In fact, any LP has a dual, obtained in a similar fasion as above. Table 1 shows the dual optimizations for a number of primal forms. Even more generally, you can create the dual for any optimization, which does not immediately fit into the standard forms, through the following set of rules:

1. Identify if the optimization is minimization or maximization. In the next steps, the instructions within the parenthesis are for maximization.
2. For any constraint, say the i^{th} constraint, introduce a dual variable u_i.
3. If the optimization is minimization (maximization), the dual optimization is a maximization (minimization), which gives a lower (upper) bound of the primal optimization. In the dual score (cost) function, the coefficient of u_i is the constant term b_i of the i^{th} constraint.
4. If the optimization is minimization (maximization), introduce the constraint $u_i \geq 0$ ($u_i \leq 0$) for every constraint of the form $g_i(\mathbf{x}) \geq b_i$ and $u_i \leq 0$ ($u_i \geq 0$) for every constraint of the form $g_i(\mathbf{x}) \leq b_i$. For any equality constraint leave its corresponding variable u_i unconstrained.
5. for every primal variable x_j, introduce an equality constraint, where the coefficient of u_i equals the coefficient of x_j in the i^{th} constraint. The constant is c_j, the coefficient of x_j in the cost.

Example 11. Inequality constraints in the dual form: Let us obtain the third row in Table 1. It shows how inequality constraints (other than $\mathbf{u} \leq \mathbf{0}$) can appear in the dual optimization. Consider the left hand side optimization of the third row. For the constraints $\mathbf{A}\mathbf{x} \leq \mathbf{b}$, we have dual variables $\mathbf{u} \leq \mathbf{0}$. For the constraints $\mathbf{x} \geq \mathbf{0}$ take dual parameters $\mathbf{s} \geq 0$. Then, the dual optimization by the above rule is given by

$$
\begin{aligned}
&\max_{\mathbf{u},\mathbf{s}} \mathbf{b}^T\mathbf{u} \\
&\text{subject to} \\
&\mathbf{A}^T\mathbf{u} + \mathbf{s} = \mathbf{c} \\
&\mathbf{u} \le \mathbf{0} \\
&\mathbf{s} \ge \mathbf{0}
\end{aligned} \tag{44}
$$

It is observed that the variables **s** are slack variables in the dual optimization. Hence, they can be removed to obtain

$$
\begin{aligned}
&\max_{\mathbf{u}} \mathbf{b}^T\mathbf{u} \\
&\text{subject to} \\
&\mathbf{A}^T\mathbf{u} \le \mathbf{c} \\
&\mathbf{u} \le \mathbf{0}
\end{aligned} \tag{45}
$$

As a general rule, the dual parameters for the positivity constraints $\mathbf{x} \ge \mathbf{0}$ are slack variables and yield to inequality constraints when they are removed.

The important result that we established here is that the optimal value of the dual optimization is always less than the optimal value of the primal one. This is called the **weak duality** theorem. If you take a closer look at the above argument, you will notice that we did not make use of the fact that the variable space is $\mathbb{R}^n$. In fact, the variable space does not need to be $\mathbb{R}^n$ in order to obtain the weak duality result. Notice that in this case, the primal is not an LP anymore. However, the variables of the dual optimization are always real and the dual is LP.

Example 12. Consider the integer optimization

$$
\begin{aligned}
&\min_{(x_1,x_2)\in\mathbb{Z}^2} -x_1 - 0.64x_2 \\
&\text{subject to} \\
&50x_1 + 31x_2 \le 250 \\
&-3x_1 + 2x_2 \le 4 \\
&x_1 \ge 0 \\
&x_2 \ge 0
\end{aligned} \tag{46}
$$

Notice that this is an integer program and not LP. Still, the dual optimization can be calculated by the third row in Table 1, which leads to

$$
\begin{aligned}
&\max_{(u_1,u_2)\in\mathbb{R}^2} 250u_1 + 4u_2 \\
&\text{subject to} \\
&50u_1 - 3u_2 \le -1 \\
&31u_1 + 2u_2 \le -0.64 \\
&u_1 \le 0 \\
&u_2 \le 0
\end{aligned} \tag{47}
$$

Solving the dual with CVX will lead to the optimal value -5.098. Hence, the minimum value in (46) is larger than -5.098. In fact, the minimum value of (46) is obtained at $x_1 = 5, x_2 = 0$, with the optimal value -5. It is said that the **duality gap** in this case equals $-5 - (-5.098) = 0.098$.

We saw that dual optimizations provide lower bounds for LP minimizations. In Example 12, we also observed that the duals can provide lower bounds for other optimizations, which are similar to LPs, but different in their variable domain. Now, let us focus again on LPs i.e., the cases where the variables are real. A more profound result in this case is that not only the dual optimal value is smaller than the primal one, but also they are *equal.* This result is called **strong duality** and is more difficult to prove. Hence, we neglect the proof in this course.

Example 13. Consider the so called **LP relaxation** of the integer optimization (46), which is obtained by modifying its variable space to $\mathbb{R}^2$:

$$\begin{aligned} \min_{(x_1,x_2)\in\mathbb{R}^2} & -x_1 - 0.64x_2 \\ & \text{subject to} \\ & 50x_1 + 31x_2 \le 250 \\ & -3x_1 + 2x_2 \le 4 \\ & x_1 \ge 0 \\ & x_2 \ge 0 \end{aligned} \tag{48}$$

Notice that the dual of (48) is exactly the one in (47), but this time the optimal value of the primal can be calculated by CVX to obtain exactly the same value as in (47):-5.098. This is obtained at $x_1 = 1.9482..., x_2 = 4.9223....$ In this case, the duality gap is zero.

One of the advantages of the CVX package is that you do not need to explicitly write the dual and separately encode in CVX. In the class, I will show you how to solve by CVX the dual program simultaneously with the primal one.