1 Duality

In the lectures, you were introduced to the concept of duality. In this recitation, we try to motivate this idea and give some examples.

1.1 Motivating Example

We start by looking at linear programs. Suppose we are trying to solve the following optimization problem:

$$\begin{aligned} \min_{x,y,z} \quad & 5x + y + 7z \\ \text{subject to} \quad & x + 2y + 3z \geq 13 \\ & -4x + 7y \geq 9 \\ & 3x - 6y - z \geq 5 \\ & x, y, z \geq 0 \end{aligned}$$

How do we approach this problem? We can try to do some algebraic manipulations on the inequalities, but we quickly see that there's no straightforward way to approach these types of problems in general. We can, however, start by trying to get some **lower bounds** on our minimization problem.

If $m = \min_{x,y,z}(5x + y + 7z)$, what can we say about m based on our constraints?

Attempt 0 A trivial lower bound would be **0**. We know that $x, y, z \geq 0$, so we can be certain that $5x + y + 7z \geq 0$. This is by no means a good lower bound: $5x + y + 7z$ is almost surely greater than 0, but it's a good place to start.

Attempt 1 If we look at the third expression, we see that $3x - 6y - z \geq 5$. Because $x, y, z \geq 0$, then $5x + y + 7z \geq 3x - 6y - z \geq \mathbf{5}$. We now have a better lower bound to our problem!

Attempt 2 Using both the first and third expressions, we get that $2(x + 2y + 3z) + 1(3x - 6y - z) \geq 2(13) + 1(5)$, or equivalently, $5x - 2y + 5z \geq 31$. Since $5x + y + 7z \geq 5x - 2y + 5z \geq \mathbf{31}$, so we now have an even better lower bound for $5x + y + 7z$.

Attempt 3 Now we use all 3 expressions. Observe that $5x + y + 7z = 3(x + 2y + 3z) + 1(-4x + 7y) + 2(3x - 6y - z)$, so we get $5x + y + 7z \geq 3(13) + 1(9) + 2(5) = \mathbf{58}$. This gives us a really good lower bound!

Hidden behind these equations, of course, is the true meat of the problem – how did we come up with these numbers to magically add up to our objective function? This is what we will explore next:

1. How do we get the "best" lower bounds for our minimization problem?
2. Is 58 truly the best lower bound? Could there exist x, y, z such that $5x + y + 7z$ is larger than 58?

1.2 Dual Problem

In this section, we attempt to solve the first question posed above using standard algebraic tools.

Recall our optimization problem:

$$\begin{aligned}\min_{x,y,z} \quad & 5x + y + 7z \\ \text{subject to} \quad & x + 2y + 3z \geq 13 \\ & -4x + 7y \geq 9 \\ & 3x - 6y - z \geq 5 \\ & x, y, z \geq 0\end{aligned}$$

In the previous section, we were multiplying our constraints by constants. Let the constants we multiply to the first, second, and third equations be a, b, c respectively.

$$\begin{aligned}\min_{x,y,z} \quad & 5x + y + 7z \\ \text{subject to} \quad & a(x + 2y + 3z) \geq 13a \\ & b(-4x + 7y) \geq 9b \\ & c(3x - 6y - z) \geq 5c \\ & x, y, z \geq 0\end{aligned}$$

For the signs to stay the same, we need to have $a, b, c \geq 0$. Adding the three equations, we get

$$(a - 4b + 3c)x + (2a + 7b - 6c)y + (3a - c)z \geq 13a + 9b + 5c$$

We want this to be a **lower bound** for our objective function, i.e.

$$5x + y + 7z \geq (a - 4b + 3c)x + (2a + 7b - 6c)y + (3a - c)z \geq 13a + 9b + 5c$$

To get the **best possible lower bound**, we thus solve

$$\begin{aligned}\max_{a,b,c} \quad & 13a + 9b + 5c \\ \text{subject to} \quad & a - 4b + 3c \leq 5 \\ & 2a + 7b - 6c \leq 1 \\ & 3a - c \leq 7 \\ & a, b, c \geq 0\end{aligned}$$

This is called our **dual** problem.

The dual might NOT necessarily be easier to solve than the primal (as we can observe in our example). It can be helpful (as we will observe later on in the section on KKT conditions), but in our current example it is not. The key purpose of re-expressing our problem in this dual form is to demonstrate the following ideas:

1. We re-express our **minimization problem** as a **maximization problem**.
2. We introduce **new variables which correspond to the constraints** of our original problem.
3. The new (dual) problem optimizes over these new variables and gives a bound on the solution of the original (primal) problem.

1.2.1 Exercise 1

Rewrite the primal and dual problems in the previous page using dot products and matrices.

1.2.2 Exercise 2

Show that the dual of the dual is the primal.

1.2.3 Exercise 3

Write the dual of the following optimization problem. (How do we deal with the = sign?)

$$\begin{aligned} \min_{x,y,z} \quad & 3x + 2y + 4z \\ \text{subject to} \quad & 2x + 5y + z \leq 14 \\ & 4x + 3y - 6z = -3 \\ & -x + 4y - 2z \geq 4 \\ & x, y, z \geq 0 \end{aligned}$$

1.3 General Form

Primal

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

Dual

$$\begin{aligned} \max_y \quad & b^\top y \\ \text{subject to} \quad & A^\top y \geq c \\ & y \geq 0 \end{aligned}$$

There exist other forms to write the primal and dual. Some can easily be derived using tricks we've previously seen, while others take a bit more work. More details can be found here. [1]

1.4 Weak and Strong Duality Theorems

Weak Duality Theorem. *In the above formulation,* $\max_y b^\top y \leq \max_x c^\top x$.

Pf. $c^\top x = x^\top c \leq x^\top (A^\top y) = (Ax)^\top y \leq b^\top y$.

The weak duality theorem formalizes the notion we had earlier of the dual solution being a lower bound of the primal solution.

Recall that we had one more unanswered question from page 1: "Is 58 truly the best lower bound? Could there exist x, y, z such that $5x + y + 7z$ is larger than 58?" In other words, when is the dual linear bound tight? To answer this, we need strong duality.

Strong Duality *Under certain conditions,* $\max_y b^\top y = \max_x c^\top x$.

One such condition where the strong duality holds is if the primal is a feasible convex objective with linear constraints. This is the case for most of the problems we've seen so far. (We won't be proving strong duality.)

[1] https://en.wikipedia.org/wiki/Dual_linear_program#Vector_formulations

2 Lagrange Dual

Not all problems can be written as linear programs. Not all objective functions are of the form $c^\top \mathbf{x}$. For example, solving the SVM is a quadratic programming problem. How do we deal with more general objective functions? Recall our motivation for using the dual in the previous section:

1. We re-express our **minimization problem** as a **maximization problem**.
2. We introduce **new variables which correspond to the constraints** of our original problem.
3. The new (dual) problem optimizes over these new variables and gives a bound on the solution of the original (primal) problem.

It turns out that we can apply these same principles to more general problems!

2.1 Lagrange Dual Formulation

Given a primal problem

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & h_i(x) \leq 0, \quad i = 1, \dots, m \\ & l_j(x) = 0, \quad j = 1, \dots, r \end{aligned}$$

We define the Lagrangian as

$$L(x, u, v) = f(x) + \sum_{i=1}^{m} u_i h_i(x) + \sum_{j=1}^{r} v_j l_j(x), \qquad u_i \geq 0$$

$L(x, u, v)$ introduces new variables which correspond to our original constraints.
Note that for any $u_i \geq 0$ and v_j, $f(x) \geq L(x, u, v)$.

In fact, $f(x) = \max_{u \geq 0, v} L(x, u, v)$.
We hence express our original optimization (**primal**) problem as

$$\min_{x} f(x) = \boxed{\min_{x} \max_{u \geq 0, v} L(x, u, v)}$$

Our **dual** problem, meanwhile, is expressed as

$$\boxed{\max_{u \geq 0, v} \min_{x} L(x, u, v)} = \max_{u \geq 0, v} g(u, v)$$

As seen above, $g(u, v)$ is defined as $\min_x L(x, u, v)$.

The dual problem can also equivalently be expressed as:

$$\begin{aligned} \max_{u, v} \quad & g(u, v) \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

2.2 Weak and Strong Duality

Let f^* be a solution to our primal problem $\min_x f(x)$ and g^* be a solution to our dual problem $\max_{u \geq 0, v} g(u, v)$.

Weak duality For any feasible x and feasible u, v,

$$f^* \geq g^*$$

Strong duality We say strong duality holds if

$$f^* = g^*$$

The weak duality theorem formalizes the notion that the dual objective function is a **lower bound** of the primal objective function. When strong duality holds, then this bound is tight.

Slater's condition If the primal is a convex problem, i.e. f and h_i are convex, and l_i are affine, and there exists at least one strictly feasible point x, meaning

$$h_i(x) < 0 \ \ \forall i \ \ \textbf{and} \ \ l_j(x) = 0 \ \ \forall j$$

then strong duality holds.

Convexity The dual problem is always convex (or concave), even though the original primal problem is non-convex.

3 KKT conditions

For a problem with strong duality, x^* and u^*, v^* are primal and dual solutions if and only if x^* and u^*, v^* satisfy the Karush-Kuhn-Tucker (KKT) conditions:

- **Stationarity:** $0 \in \partial(L(x, u, v))$
- **Complementary slackness:** $u_i \cdot h_i(x) = 0$ for all i
- **Primal Feasibility:** $h_i(x) \leq 0$ and $l_j(x) = 0$ for all i, j
- **Dual Feasibility:** $u_i \geq 0$ for all i

Here, ∂ represents the subgradient. If all terms in L are differentiable, then the stationarity condition can be expressed as $\nabla L(x, u, v) = \nabla f(x) + \sum_{i=1}^{m} u_i \nabla h_i(x) + \sum_{j=1}^{r} v_j \nabla l_j(x) = 0$

Example 1 Use KKT conditions to solve the following constrained optimization problem.

$$\begin{aligned} \min_{x,y} \quad & \frac{1}{2}x^2 + \frac{1}{2}y^2 - xy - 3y \\ \text{subject to} \quad & x + y = 3 \end{aligned}$$