UNIVERSITY OF CALIFORNIA, BERKELEY

Lecture 14 - EE127 Different Interpretations of Duality

Oscar Ortega

July 16, 2021

1 EXAMPLE: LINEAR PROGRAM

Consider the following problem:

$$\min c^T x$$
$$s.t: Ax \leq b$$

Note: Recall that the inequality when comparing vectors is considered as a component-wise operation, i.e $x \leq y \rightarrow x_i \leq y_i : i = 1, 2, ..., n$.
Given the definition of the Lagrangian:

$$\mathcal{L}(x, \lambda, \nu) := f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x) + \sum_{i=1}^{n} \nu_i h_i(x)$$
$$\lambda_i \geq 0 : i = 1, 2, ..., m$$

One can see the Lagrangian is as follows:

$$\mathcal{L}(x, \lambda) = c^T x + \lambda^T (Ax - b)$$

Note how there are not equality constraints.

$$= c^T x + \lambda^T Ax - \lambda^T b$$
$$= -\lambda^T b + (\lambda^T A + c) x$$

Recall how we define the function g:

$$g(\lambda) = \min_x \mathscr{L}(x, \lambda)$$

Which in this example is equal to the following:

$$= \min_x (-b^T a + A^T \lambda + c)x$$

$$= \begin{cases} -\infty & A^T \lambda + c \neq 0 \\ -b^T \lambda & \text{otherwise} \end{cases}$$

Because our function g is a lower bound on the value of min of the primal problem, we then want to maximize lambda.
This results in the following problem:

$$\max_\lambda g(\lambda)$$

$$\max_\lambda - b^T \lambda$$

$$A^T \lambda + c = 0$$

$$\lambda \geq 0$$

Where the following formulation is known as the **dual problem**, as opposed to the **primal problem** we were originally presented with.

2 Shadow-Prices Interpretation

If we consider a general optimization problem:

$$\min f_0(x)$$

$$\text{s.t } f_i(x) \leq 0 : i = 1, 2, ..., n$$

Lagrange duality gives us an interesting economic interpretation. Suppose the variable x is how our enterprise ('go back to the wine-making example'). We can interpret our optimization as a minimization of the cost to run this company, subject to some resource constraints to the way we run our company(our $f_i's$). This would mean our optimal cost to run our company is p^* and our optimal profit is the negative cost $= -p^*$
So, what if we no longer had to strictly follow the resource constraints and instead had to pay λ_i cost per unit to violate the resource $f_i(x)$ or where given λ_i cost per unit to under-use a given resource, i.e $f_i(x) < 0$, Well then this would be optimizing the following objective function.

$$\mathscr{L}(\lambda, x) : \lambda \geq 0 = f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x)$$

We can think of the minimization of x as maximizing the price over the worst case penalty/'prices', i.e it would never be better to overuse a resource. This would yield an optimal price d^* subject to this dual form where the λ^* is the set of prices for which there for which there is no advantage to pay violations or to under-use a resource, these are the **shadow prices** of our original problem.

3 Interpretation 2: Mini-max Theorem

Mini-max Theorem:
For any sets X, Y and any function $F : (X, Y) \to \mathbf{R}$:

$$\max_y \min_x F(x, y) \le \min_x \max_y F(x, y)$$

Proof:
Fix: (x_0, y_0) and lets define $h(y_0) = \min_{x\in X} F(x, y_0)$

$$\le F(x_0, y_0) \le \max_{y\in Y} F(x_0, y) = g(x_0)$$

$$\forall (x, y) : h(y_0) \le g(x_0)$$

$$\max_{x\in Y} h(y) \le \min_{x\in X} g(x)$$

$$\to \max_y \min_x F(x, y) \le \min_x \max_y F(x, y)$$

With this in mind, let us consider the following optimization problem:

$$p^* = \min f_0(x)$$

$$\text{s.t} f_i(x) \le 0$$

and the following function:

$$\sup_{\lambda\ge 0}(f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x))$$

$$= \begin{cases} f_0(x) & f_i(x) \le 0 : i = 1, 2, ..., n \\ \infty & \text{otherwise} \end{cases}$$

Note how if we chose x with some constraint greater $f_i(x) \ge 0$), then choosing a $\lambda \to \infty$ would make the value of this function jump to ∞ as well. This means we can alternatively define p^* as follows.

$$p^* = \inf_x \sup_{\lambda\ge 0} L(x, \lambda)$$

If we then define d^* as the optimal value for the dual problem.

$$d^* = \sup_{\lambda\ge 0} \inf_x L(x, \lambda)$$

We can immediately conclude $d^* \le p^*$ from the mini-max theorem and view weak duality from this new standpoint. We can also view strong duality as when equality is obtained in the given equation.

3.1 Game Interpretation

We can interpret the mini-max theorem in terms of a continuous zero-sum game in the sense both players have infinite choices. For our purposes, let player one give action's w try to minimize $f(w,z)$, and let player two give action's z to maximize $f(w,z)$. Assuming player one went first and player two went after learning player one's response. Our score would be the following after one round of play.

$$\min_z \max_w F(z,w)$$

Now, if we let player two choose first and have player one choose after learning player two's actions, our score would be as follows.

$$\max_w \min_z F(z,w)$$

We know intuitively that it is better for player two to go after learning from player one because he no longer has to prepare for player one's worst case action. Assuming this is true we can arrive at the following.

$$\min_z \max_w F(z,w) \leq \max_w \min_z F(z,w)$$

This is the just the minimax inequality.

4 Interpretation 3: Geometry-Based

Consider the following minimization:

$$\min f_0(x)$$

$$\text{s.t} f_i(x) \leq 0$$

And keep the following example in mind:

$$\min x^2$$

$$x \leq 3$$

Let us define the following reparameterization:

$$\min t$$

$$(u,t) \in G$$

Where we define G as follows:

$$G = \{(f_1(x), f_0(x)) \in \mathbf{R}^2, f_1(x) \leq 0\}$$

$$p^* = \inf\{t : (u,t) \in G, u \leq 0\}$$

In terms of the example above:

$$G = \{(x+3, x^2) : x \in \mathbf{R}\, x+3 \leq 0$$

With this in mind, we can define the Lagrangian for the example as follows:

$$\mathcal{L}(u,t,\lambda) = t + \lambda u$$

$$\lambda \geq 0$$

$$(u,t) \in G$$

This implies that we can reformulate the optimization problem as the following:

$$= \min_{(u,t)\in G} t + \lambda u, \lambda \geq 0$$

To motivate the following, consider the following minimization of our dual for a **fixed** λ

fixed.JPG

As we can see, the minimum of our dual for a fixed lambda, 'the y intercept on the graph', will always be achieved if we have a hyperplane **tangent** to the convex set.
Now, lets try to minimize the following function in terms of λ.

$$g(\lambda) = \inf_{(u,t)\in G,\lambda} t + \lambda u$$

Given the graph of the function and given that we are trying to minimize this function in terms of λ what can we immediately see? The existence of solutions in the interior of the set, implies that we cannot immediately minimize this function to $-\infty$. In other words, the fact that some of the valid intercepts in the epigraph of our curve tells us we cannot set the slope of λ to $-\infty$ and achieve a minimum intercept $-\infty$ as well. Furthermore, this is sufficient to ensure strong duality if the function is convex! This motivates what is known as **Slater's condition**.

5 Connecting this to the notion of strong duality

5.1 Slater's condition for Strong Duality

Slater's condition is the following:

$$\exists x \in \text{domain s.t } f_i(x) < 0, h_i(x) = 0$$

We then know if this condition is true, then strong duality will hold and in the case of a convex problem, $p^* = d^*$, where the following are the optimal values of the primal and dual formulations of a convex optimization problem. We call points that satisfy these conditions **Strictly Feasible**.