UNIVERSITY OF CALIFORNIA, BERKELEY

Lecture 13: EE127 - Convexity 3 - Duality

Oscar Ortega

July 16, 2021

0.1 Relaxation Proof

Consider the following optimization problem:

$$p^* = \min_{x \in \mathscr{X}} f_0(x)$$
$$b(x) = \mu$$

$$g^* = \min_{x \in \mathscr{X}} f_0(x)$$
$$b(x) \leq \mu$$

Recall from the previous lecture that if the following hold, we could state that when replacing an equality constraint. Equality of the two optimums would occur if the following held:

- the objective function-increasing over X
- b is non-decreasing over x
- the optimal p^* is attained at some optimal point x^*, and the optimal value g^* is attained at some optimal point $\tilde{x}$

Let us assume that $g^* < p^*$ for the sake of contradiction

1. we know $u = b(x^*)$ because u is a feasible point of the first program and $u \leq b(\tilde{x})$ by a similar reasoning
2. $b(\tilde{x}) < u = b(x^*)$

3. $\tilde{x} \leq x^*$ because b is monotonically increasing

4. $f_0(\tilde{x}) \geq f_0(x^*)$ because f is monotonically decreasing

5. $g^* = f_0(x^*) \geq f_0(x^*) = p^*$

6. $g^* \geq p*$

Therefore, they must be equal.

1 Duality

consider our standard optimization problem:
minimize $f_0(x)$
subject to $f_i(x) \leq 0 \forall i : 1 \leq i \leq m$
$h_i(x) = 0 : i : 1 \leq i \leq m$

How do we we write an algorithm to figure out what the optimal point is?
We define the **Lagrangian** as follows:

$$\mathscr{L}(x, \lambda, v) = f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x) + \sum_{i=1}^{n} v_i h_i(x)$$

$$\lambda_i \geq 0$$

λ_i, v_i, are called **Lagrange Multipliers**, and also known as **dual variables**.

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

things to note:

- g is a function of λ, μ
- $\mathscr{L}$ is an affine function of λ, μ
- $g(\lambda, v)$ is concave

We can also provide a lower bound on the optimal value * based on the function g:
Proof:
Let $\tilde{x}$ a feasible point of the set.

$$f_i(x) \leq 0 \rightarrow \sum_i \lambda_i f_i(x) \leq 0$$

$$h_i(x) = 0 \rightarrow \sum_i v_i h_i(x) = 0$$

$$g(\mu, v) \leq f_0(x) : \forall x$$

We might not be always be able to compute the minimum of an objective function, but we know the Lagrangian will always give you the maximum of a function.
Where is this coming from? consider the following minimization problem:
Let

$$\mathbf{1}_{neg}(u) = \{0 \text{ if } u \leq 0, \infty \text{ otherwise}\}$$

Let

$$\mathbf{1}_0(u) = \{0 \text{ if } u = 0, \infty \text{ otherwise}\}$$

$$\min f_0(x) + \sum_i \mathbf{1}_{\mathbf{neg}} f_i(x) + \sum_i \mathbf{1}_0(h_i(x))$$

If we violate a constraint in the following program, this is really really bad for us, so we can consider the Lagrangian as a softer version of the following function.

1.1 Example: Minimum Norm

Allow A to be full row rank, and allow A to be lower determined.

$$\min x^T x$$

$$\text{s.t } Ax = b$$

$$\mathscr{L}(x, v) = x^T x + v^T (Ax - b)$$

$$g(v) = min_x L(x, v)$$

$$\nabla x \mathscr{L}(x, v) = 2x + A^T v$$

$$x = -\frac{1}{2} A^T v$$

This is a unique minimizer because $x^T x$ is an affine transformation of a PSD matrix.

$$g(v) = \frac{1}{4} v^T AA^T v + v^T (-\frac{1}{2} AA^T v - b)$$

$$= -\frac{1}{4} v^T AA^T v - v^T b \leq \min\{x^T x | Ax = b\}$$

Our natural tendency therefore is to try and maximize this lower bound.

2 Lagrange Dual Problem

We can now define the **Lagrangian dual formulation** as follows

$$\max g(\lambda, v) \text{ s.t } \lambda \geq 0$$

Applying this to the previous problem:

$$g(v) = -\frac{1}{4} v^T AA^T v - v^T b$$

$$\nabla_v g(v) = -\frac{1}{2}(AA^T)v - b$$
$$v^* = -2(AA^T)^{-1}b$$
$$g(v^*) = A^T(AA^T)^{-1}b$$

How did we know that the two values would coincide?
Definitions:
weak duality:$d^* \leq p^*$
strong duality: $p^* = d^*$
We will learn about when strong duality versus weak duality holds later in this course.

3 Interpretations of Duality

3.1 wine-making

Lets say, we have 2000kg or Merlot, and 1500 kg of Shiraz and let us consider two blends of of the grapes to make a bottle of wine.

- 10 m, 4 s: 20
- 8m, 4s: 15
- 7m, 15s: 25

Our goal is to want to maximize this money.

$$\max(20b_1 + 15b_2 + 25b_3)$$
$$b_1, b_2, b_3 \geq 0$$
$$10b_1 + 8b_2 + 7b_3 \leq 2000$$
$$5b_1 + 4b_2 + 15q_2 \leq 1500$$

In general: if we have a program of the following form,

$$\min c^T x \iff \max -b^T A$$
$$\text{s.t } Ax \leq b \iff \text{s.t } A^T + c = 0, \lambda \geq 0$$

This would make the dual formulation as follows:

$$\min(2000\lambda_1 + 1500\lambda_2)$$
$$\text{s.t } 10\lambda_1 + 5\lambda_2 \geq 20$$
$$8\lambda_1 + 4\lambda_2 \geq 15$$
$$7\lambda_1 + 15\lambda_2 \geq 25$$

We can interpret this as trying to minimize the amount of grapes left over subject to the gain per bottle being greater than the price of the bottle.