DUALITY IN CONVEX OPTIMIZATION

CHRISTOPHER J. HILLAR

1. Introduction

Consider the problem [1]:

$$\text{minimize } f_0(x)$$
$$\text{subject to } f_i(x) \leq 0$$
$$h_i(x) = 0.$$

where $f_0(x)$ is a convex function and the $h_i(x)$ are affine. Set $\mathcal{D}$ to be the intersection of the domains of the f_i. We set p^* to be the solution to this minimization problem.

Define the *Lagrangian* associated to the problem to be

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x).$$

The λ and ν are called *dual variables.*

We would like to minimize this function over the domain:

$$g(\lambda, \nu) := \inf_{x \in \mathcal{D}} L(x, \lambda, \nu).$$

We define the *domain* of g to be those points for which g is not $-\infty$:

$$\text{dom } g = \{(\lambda, \nu) : g(\lambda, \nu) > -\infty\}.$$

Theorem 1.1. *When $\lambda \geq 0$, we have*

$$g(\lambda, \nu) \leq p^*.$$

Proof. Clearly, $g(\lambda, \nu) \leq L(x, \lambda, \nu)$ for all $x \in \mathcal{D}$. In particular, for feasible $\tilde{x}$, we have

$$g(\lambda, \nu) \leq f_0(\tilde{x}),$$

since λ are all nonnegative and $h_i(\tilde{x}) = 0$. □

It follows that we can obtain a lower bound on p^* by solving the following optimization problem, called the *Lagrange dual problem* associated to the original optimization problem:

$$\text{maximize } g(\lambda, \nu)$$
$$\text{subject to } \lambda \geq 0.$$

We will denote the answer to this problem as d^*. Of course, we have

$$d^* \leq p^*,$$

a property we call *weak duality.* This turns out to be a convex optimization problem since the objective to be maximized is concave and the constraint is convex. This concavity can be proved directly using the fact that the infimum of a sum is greater than or equal to the sum of the infimums.

2. Strong Duality

Under mild conditions, we show that both optimization problems have the same solution; that is, $d^* = p^*$.

To begin, we define the set

$$\mathcal{A} = \{(u, v, t) : \exists x \in \mathcal{D},\ f_i(x) \le u_i,\ i = 1, \dots, m,\ h_i(x) = v_i,\ i = 1,\ \dots, p,\ f_0(x) \le t\},$$

which is convex since each of the $f_i(x)$ are convex functions.

The optimal solution to our original problem is:

$$p^* = \inf \{t : (0, 0, t) \in \mathcal{A}\}.$$

Next, notice that if $\lambda \ge 0$,

$$g(\lambda, \nu) = \inf \{(u, v, t)(\lambda, \nu, 1)^T : (u, v, t) \in \mathcal{A}\}.$$

If $\lambda \ge 0$ and ν are given and $g(\lambda, \nu)$ is finite, then

$$(u, v, t)(\lambda, \nu, 1)^T \ge g(\lambda, \nu) \tag{2.1}$$

defines a (nonvertical) *supporting hyperplane* (really, half-space) to $\mathcal{A}$.

In particular, since $(0, 0, p^*) \in \operatorname{bd} \mathcal{A}$, we have

$$p^* = (\lambda, \nu, 1)^T(0, 0, p^*) \ge g(\lambda, \nu).$$

To rephrase in this language, Strong duality holds if and only if we have equality in equation (2.1).

3. Slater's constraint qualification

Definition 3.1 (Slater's Condition)**.** *There exists an $\tilde{x} \in$ relint $\mathcal{D}$ with $f_i(\tilde{x}) < 0$ for $i = 1, \dots, m$ and $A\tilde{x} = b$.*

In other words, there exists a *strictly feasible point.*

Theorem 3.2. *Slater's condition implies strong duality.*

Proof sketch. Assume for simplicity that relint $\mathcal{D}$ = int $\mathcal{D}$ and that A has full rank p. Consider the following (convex) set

$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times R : s < p^*\},$$

which is obviously disjoint from $\mathcal{A}$.

By the separating hyperplane theorem, there exists $(\tilde{\lambda}, \tilde{\nu}, \mu) \ne 0$ and α such that

$$(u, v, t) \in \mathcal{A} \Rightarrow u\tilde{\lambda}^T + v\tilde{\nu}^T + t\mu \ge \alpha,$$

and

$$(u, v, t) \in \mathcal{B} \Rightarrow u\tilde{\lambda}^T + v\tilde{\nu}^T + t\mu \le \alpha,$$

This implies that $\tilde{\lambda} \ge 0$, by the first equation (since $\mathcal{A}$ is closed under u getting larger), and $\mu \ge 0$, by the second, which says that $\mu t \le \alpha$ for all $t < p^*$ and thus $\mu p^* \le \alpha$.

For any $x \in \mathcal{D}$, we have

$$(f_1(x), \ldots, f_m(x), h_1(x), \ldots, h_p(x), f_0(x)) \in \mathcal{A},$$

and therefore by the first inequality, it follows that

$$\sum_{i=1}^{m} \tilde{\lambda}_i f_i(x) + \tilde{\nu}(Ax - b) + \mu f_0(x) \geq \alpha \geq \mu p^*. \tag{3.1}$$

We proceed in two cases.

Case 1: $\mu > 0$. Dividing (3.1) by μ, we obtain

$$L(x, \tilde{\lambda}/\mu, \tilde{\nu}/\mu) \geq p^*,$$

for all $x \in \mathcal{D}$. Thus, minimizing over x, it follows that $g(\lambda, \nu) \geq p^*$ where $\lambda = \tilde{\lambda}/\mu$ and $\nu = \tilde{\nu}/\mu$. By weak duality, we have $g(\lambda, \nu) = p^*$.

Case 1: $\mu = 0$. Using (3.1), it follows that for $\tilde{x}$ satisfying Slater's condition, we have

$$\sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{x}) \geq 0,$$

and therefore $\tilde{\lambda} = 0$ since all $f_i(\tilde{x}) < 0$ and $\tilde{\lambda} \geq 0$. From $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ and $\tilde{\lambda} = \mu = 0$, we conclude that $\tilde{\nu} \neq 0$. Thus, (3.1) implies that

$$\tilde{\nu}(Ax - b) \geq 0.$$

By assumption, $\tilde{\nu}(A\tilde{x} - b) = 0$, and since $\tilde{x} \in \text{int } \mathcal{D}$, it follows that there exists a perturbation $x \in \mathcal{D}$ such that $\tilde{\nu}(Ax - b) < 0$ unless $\tilde{\nu}A = 0$. This contradicts the fact that $\text{rank}(A) = p$. $\square$

References

[1] S. Boyd, L. Vandenberghe, *Convex Optimization*.