1 Duality Theory for Constrained Optimization Problems

To begin our investigation, consider the following primal problem:

$$(P) \qquad \begin{array}{rccll} v_p^* & = & \inf & f(x) & \\ & & \text{subject to} & g_i(x) \le 0 & \text{for } i = 1, \ldots, m_1, \\ & & & h_j(x) = 0 & \text{for } j = 1, \ldots, m_2, \\ & & & x \in X. & \end{array}$$

Here, $f, g_1, \ldots, g_{m_1}, h_1, \ldots, h_{m_2} : \mathbb{R}^n \to \mathbb{R}$ are *arbitrary* functions, and X is an *arbitrary* non–empty subset of $\mathbb{R}^n$. For the sake of brevity, we shall write the first two sets of constraints in (P) as $g(x) \le \mathbf{0}$ and $h(x) = \mathbf{0}$, where $g : \mathbb{R}^n \to \mathbb{R}^{m_1}$ is given by $g(x) = (g_1(x), \ldots, g_{m_1}(x))$ and $h : \mathbb{R}^n \to \mathbb{R}^{m_2}$ is given by $h(x) = (h_1(x), \ldots, h_{m_2}(x))$.

Now, the **Lagrangian dual problem** associated with (P) is the following problem:

$$(D) \qquad \begin{array}{rccl} v_d^* & = & \sup & \theta(u, v) \equiv \inf_{x \in X} L(x, u, v) \\ & & \text{subject to} & u \in \mathbb{R}_+^{m_1},\ v \in \mathbb{R}^{m_2}. \end{array}$$

Here, $L : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \to \mathbb{R}$ is the Lagrangian function given by

$$L(x, u, v) = f(x) + \sum_{i=1}^{m_1} u_i g_i(x) + \sum_{j=1}^{m_2} v_j h_j(x) = f(x) + u^T g(x) + v^T h(x). \tag{25}$$

Observe that the above formulation is reminiscent of the penalty function approach, in the sense that we incorporate the primal constraints $g(x) \le \mathbf{0}$ and $h(x) = \mathbf{0}$ into the objective function of (D) using the Lagrange multipliers u and v. Also, since the set X is arbitrary, there can be many different Lagrangian dual problems for the same primal problem, depending on which constraints are handled as $g(x) \le \mathbf{0}$ and $h(x) = \mathbf{0}$, and which constraints are treated by X. However, different choices of the Lagrangian dual problem will in general lead to different outcomes, both in terms of the dual optimal value as well as the computational efforts required to solve the dual problem.

Let us now investigate the relationship between (P) and (D). For any $\bar{x} \in X$ and $(\bar{u}, \bar{v}) \in \mathbb{R}_+^{m_1} \times \mathbb{R}^{m_2}$, we have

$$\inf_{x \in X} L(x, \bar{u}, \bar{v}) \le f(\bar{x}) + \bar{u}^T g(\bar{x}) + \bar{v}^T h(\bar{x}) \le \sup_{u \ge \mathbf{0},\, v \in \mathbb{R}^{m_2}} L(\bar{x}, u, v). \tag{26}$$

This implies that

$$\sup_{u \ge \mathbf{0},\, v \in \mathbb{R}^{m_2}} \inf_{x \in X} L(x, u, v) \le \inf_{x \in X} \sup_{u \ge \mathbf{0},\, v \in \mathbb{R}^{m_2}} L(x, u, v). \tag{27}$$

Theorem 6 (Weak Duality Theorem) *Let $\bar{x}$ be feasible for (P) and $(\bar{u}, \bar{v})$ be feasible for (D). Then, we have $\theta(\bar{u}, \bar{v}) \le f(\bar{x})$. In particular, if $v_d^* = +\infty$, then (P) has no feasible solution.*

Example 2 *Consider the following problem from [1, Example 6.2.2]:*

$$\begin{array}{ll} \text{minimize} & f(x) \equiv -2x_1 + x_2 \\ \text{subject to} & h(x) \equiv x_1 + x_2 - 3 = 0, \\ & x \in X, \end{array} \tag{28}$$

where $X \subset \mathbb{R}^2$ is the following discrete *set:*

$$X = \{(0,0), (0,4), (4,4), (4,0), (1,2), (2,1)\}.$$

By enumeration, we see that the optimal value of (28) is -3, attained at the point $(x_1, x_2) = (2,1)$. Now, one can verify that the Lagrangian function is given by

$$\begin{aligned} \theta(v) &= \min_{x \in X} \{-2x_1 + x_2 + v(x_1 + x_2 - 3)\} \\ &= \begin{cases} -4 + 5v & \textit{for } v \leq -1, \\ -8 + v & \textit{for } -1 \leq v \leq 2, \\ -3v & \textit{for } v \geq 2. \end{cases} \end{aligned}$$

It follows that $\max_v \theta(v) = -6$, *which is attained at $v = 2$. Note that the duality gap in this example is $\Delta = -3 - (-6) = 3 > 0$.*

The above example raises the important question of when the duality gap is zero. It turns out that there is a relatively simple answer to this question. Before we proceed, let us introduce the following definition:

Theorem 7 (Strong Duality Theorem): Let L be the Lagrangian function defined in (25). Suppose that:

1. *X is a compact convex subset of $\mathbb{R}^n$,*
2. *$(u,v) \mapsto L(x,u,v)$ is continuous and concave on $\mathbb{R}_+^{m_1} \times \mathbb{R}^{m_2}$ for each $x \in X$, and*
3. *$x \mapsto L(x,u,v)$ is continuous and convex on X for each $(u,v) \in \mathbb{R}_+^{m_1} \times \mathbb{R}^{m_2}$.*

Then, we have

$$\sup_{u \geq \mathbf{0},\, v \in \mathbb{R}^{m_2}} \min_{x \in X} L(x,u,v) = \min_{x \in X} \sup_{u \geq \mathbf{0},\, v \in \mathbb{R}^{m_2}} L(x,u,v). \tag{29}$$

Example 3 (Linear Programming) *Consider the following standard form LP:*

$$\begin{array}{ll} \text{minimize} & f(x) \equiv c^T x \\ \text{subject to} & h(x) \equiv b - Ax = \mathbf{0}, \\ & g(x) \equiv -x \leq \mathbf{0}, \end{array} \tag{30}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$ are given. The Lagrangian dual associated with (30) is

$$\begin{array}{ll} \text{maximize} & \theta(u, v) \equiv \inf_{x \in \mathbb{R}^n} \left\{ c^T x - u^T x + v^T (b - Ax) \right\} \\ \text{subject to} & u \in \mathbb{R}^n_+,\ v \in \mathbb{R}^m. \end{array} \tag{31}$$

Note that for fixed $u \in \mathbb{R}^n_+$ and $v \in \mathbb{R}^m$, we have

$$\theta(u, v) = \begin{cases} b^T v & \text{if } A^T v + u = c, \\ -\infty & \text{otherwise.} \end{cases}$$

It follows that (31) can be written as

$$\begin{array}{ll} \text{maximize} & b^T v \\ \text{subject to} & A^T v \leq c, \end{array} \tag{32}$$

which is precisely the dual LP we defined before.