Nonlinear Programming, Part II: Duality and Geometric Multipliers

John MacLaren Walsh, Ph.D.

Consider the general family of optimization problems of the form

\[
\inf_{x \in \mathcal{X}} f(x) \tag{1}
\]

where \( \mathcal{X} \subseteq \mathbb{R}^N, f : \mathcal{X} \to \mathbb{R}, h_i : \mathcal{X} \to \mathbb{R}, i \in \{1, \ldots, I\}, g_j : \mathcal{X} \to \mathbb{R}, j \in \{1, \ldots, J\} \) are arbitrary (not necessarily differentiable or continuous) functions such that there is at least one feasible \( x \) and there is a finite optimum \( f^* \).

Throughout the rest of these notes we will refer to this problem as the primal problem.

For such a problem, and for a collection of multipliers \( \lambda \in \mathbb{R}^I \) and \( \mu \in \mathbb{R}^J, \mu \geq 0 \), define the Lagrangian function

\[
L(x, \lambda, \mu) = f(x) + \sum_{i=1}^I \lambda_i h_i(x) + \sum_{j=1}^J \mu_j g_j(x) \tag{2}
\]

Additionally, define the Lagrangian dual function as

\[
g(\lambda, \mu) = \inf_{x \in \mathcal{X}} f(x) + \sum_{i=1}^I \lambda_i h_i(x) + \sum_{j=1}^J \mu_j g_j(x) = \inf_{x \in \mathcal{X}} L(x, \lambda, \mu) \tag{3}
\]

Because it is the infimum over a family of affine functions of \( \lambda, \mu \), the Lagrangian dual function must be a concave function. (Indeed, affine functions are convex, as are their negatives, and the supremum of a family of convex functions is convex, hence the infimum of a series of affine functions is concave). Observe that the minimization to calculate the dual is carried out over all \( x \in \mathcal{X} \), rather than just those within the constraint set. For this reason, we observe that, for a fixed \( \lambda \in \mathbb{R}^I \) and \( \mu \geq 0 \),

\[
g(\lambda, \mu) = \inf_{x \in \mathcal{X}} L(x, \lambda, \mu) \tag{4}
\]

\[
\leq \inf_{x \in \mathcal{X}} \inf_{x \in \mathcal{X}} L(x, \lambda, \mu) \tag{5}
\]

\[
\leq \inf_{x \in \mathcal{X}} \left( h_i(x) = 0, i \in \{1, \ldots, I\} \right) \tag{6}
\]

\[
g(\lambda, \mu) \leq f^* \tag{7}
\]

where (6) holds because for \( x \) in the constraint region \( \mu_j g_j(x) \leq 0, j \in \{1, \ldots, J\} \) and \( \lambda_i h_i(x) = 0, i \in \{1, \ldots, I\} \).

Hence, if \( f^* \) is the optimal value of the primal problem, i.e.

\[
f^* = \inf_{x \in \mathcal{X}} f(x) \tag{8}
\]

Then, for any multipliers \( \lambda \in \mathbb{R}^I \) and \( \mu \geq 0 \), the dual underbounds the optimum of the primal problem

\[
g(\lambda, \mu) \leq f^* \tag{9}
\]

Because (9) holds for any viable multipliers \( \lambda \in \mathbb{R}^I \) and \( \mu \geq 0 \), the tightest underbound is obtained at the maximum over these variables, i.e.

\[
\sup_{\lambda \in \mathbb{R}^I, \mu \geq 0} g(\lambda, \mu) \leq f^* \tag{9}
\]

This fact is referred to as the weak duality theorem. We refer to the maximization of the Lagrangian dual function as the dual problem. This lower bound (9) is significant, for, when we can calculate the Lagrangian dual function \( g \), the dual problem is always a convex optimization problem, regardless of the primal problem. That is, the dual problem is always a maximization of a concave function \( g \) over a linearly constrained region \( \mu \geq 0 \) ... a convex optimization problem!
We call the difference between the right and left hand sides of (9), i.e. the primal optimal value minus the dual optimal value, the duality gap. Hence, when the right and left hand sides of (9) are equal, we say that there is no duality gap, and we can solve the primal problem by solving the dual. In the case of no duality gap, we call a collection of multipliers $\lambda^*, \mu^*$ attaining the optimal dual value geometric multipliers. That is, $\lambda^*, \mu^*$ are said to be geometric multipliers if

$$f^* = g(\lambda^*, \mu^*) = \max_{\lambda \in \mathbb{R}^J, \mu \geq 0} g(\lambda, \mu)$$

A collection $x^*, \lambda^*, \mu^*$ form an optimal solution geometric multiplier pair (and there is no duality gap) if and only if

$$x^* \in \arg \min_{x \in \mathcal{X}} L(x, \lambda^*, \mu^*), \quad h_i(x^*) = 0, \; i \in \{1, \ldots, I\}, \quad \mu^*_j g_j(x^*) = 0, \; g_j(x^*) \leq 0, \; j \in \{1, \ldots, J\}$$

(11)

To see why these conditions imply that there is no duality gap, observe that

$$L(x^*, \lambda^*, \mu^*) = \min_{x \in \mathcal{X}} L(x, \lambda^*, \mu^*)$$

(12)

$$= \inf_{x \in \mathcal{X} \mid h_i(x) = 0, \; g_j(x) \leq 0, \; i \in \{1, \ldots, I\}, \; j \in \{1, \ldots, J\}} L(x, \lambda^*, \mu^*)$$

(13)

$$= \inf_{x \in \mathcal{X} \mid h_i(x) = 0, \; g_j(x) \leq 0, \; i \in \{1, \ldots, I\}, \; j \in \{1, \ldots, J\}} f(x) = f^*$$

(14)

Here, the equality (12) holds by the first condition in (11), and the equality in (13) and (14) holds because $x^*$ obeys the constraints.

We can summarize the situation when geometric multipliers exist and there is no duality gap with the saddle point theorem: $\lambda^*, \mu^*$ are geometric multipliers and $x^*$ attains the optimal solution to the primal problem if and only if

$$L(x, \lambda^*, \mu^*) \geq L(x^*, \lambda^*, \mu^*) \geq L(x^*, \lambda, \mu)$$

(15)

for any $x \in \mathcal{X}$, $\lambda \in \mathbb{R}^J$, $\mu \geq 0$. That is, the Lagrangian function has a saddle point at $(x^*, \lambda^*, \mu^*)$.

Finally, we observe that if the primal problem is convex and everything is differentiable, so that $\mathcal{X} = \mathbb{R}^N$, $f$ is convex and differentiable, $h_i$ are linear for each $i \in \{1, \ldots, I\}$, and $g_j$ are convex for each $j \in \{1, \ldots, J\}$, then the Lagrangian $L(x, \lambda, \mu)$ is convex in $x$ for fixed $\lambda$ and $\mu$, and thus the KKT gradient condition

$$\nabla f(x^*) + \sum_{i=1}^I \lambda^*_i \nabla h_i(x^*) + \sum_{j=1}^J \mu^*_j \nabla g_j(x^*) = 0$$

in this case guarantees $x^* \in \arg \min_{x \in \mathcal{X}} L(x, \lambda^*, \mu^*)$

(16)

which together with the other two conditions (11) which are the other KKT conditions, guarantees no duality gap and that the associated $x^*, \lambda^*, \mu^*$ are an optimal primal solution geometric multiplier pair.

**Strong Duality**

Because the dual problem provides the most information when there is no duality gap, it is of interest to determine conditions under which there is no duality gap. Theorems proving the lack of a duality gap under certain conditions are said to prove strong duality.

One very important class of problems for which strong duality holds is the set of convex optimization problems obeying the Slater constraint qualification. This family of optimization problems are of the form

$$\min_{x \in \mathbb{R}^N} \begin{cases} a_i^T x = b_i, \; i \in \{1, \ldots, I\} \\ c_k^T x \leq d_k, \; k \in \{1, \ldots, K\} \\ g_j(x) \leq 0, \; j \in \{1, \ldots, J\} \end{cases}$$

(17)

with $f(x)$ a convex function of $x$, all of the equality constraints affine, every inequality constraint function $g_j(x)$ a convex function of $x \forall j \in \{1, \ldots, J\}$, and obey the additional Slater constraint qualification that there must be $x$ in the relative interior of the constraint such that

$$g_j(x) < 0, \; \forall j \in \{1, \ldots, J\}, \quad a_i^T x = b_i, \; \forall i \in \{1, \ldots, I\}, \quad c_k^T x \leq d_k, \; k \in \{1, \ldots, K\}$$

(18)

i.e. for which all of the (nonlinear convex) inequality constraints hold with strict inequality.

From this fact, we observe that for differentiable convex optimization problems (convex objective, linear equality constraints, and convex inequality constraints) obeying the Slater constraint qualification, the KKT conditions are both necessary and sufficient for $x^*$ to be the global solution to the primal problem and $\lambda^*, \mu^*$ to be associated geometric multipliers.