# Class 7: More results using duality theory

## 1. The relation with KKT conditions

Let us consider the following primal problem:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X, \end{array} \tag{P}$$

where

$$X := \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\}, \tag{1}$$

with $f\colon \mathbb{R}^n \to \mathbb{R}$, $g\colon \mathbb{R}^n \to \mathbb{R}^m$ and $h\colon \mathbb{R}^n \to \mathbb{R}^p$. Moreover, its dual problem is given by

$$\begin{array}{ll} \max & \omega(\lambda, \mu) \\ \text{s.t.} & (\lambda, \mu) \in D, \end{array} \tag{D}$$

where

$$D := \{(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p \mid \omega(\lambda, \mu) > -\infty\}.$$

and $\omega\colon D \to \mathbb{R}$ is defined by

$$\omega(\lambda, \mu) := \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu).$$

Now, we will show the relation between the dual variables and the Lagrange multipliers that appear in the KKT conditions of the primal problem.

**Theorem 1.** *For problem (P), assume that $f$ and $g$ are convex and that $h$ is affine. Suppose also that the Slater CQ holds for (P), i.e., there exists $\hat{x} \in \mathbb{R}^n$ such that $h(\hat{x}) = 0$ and $g_i(\hat{x}) < 0$ for all $i = 1, \ldots, m$. Then, $x^* \in X$ is an optimal solution of the primal problem (P) if, and only if, there exists $(\lambda^*, \mu^*) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ such that*

$$L(x^*, \lambda^*, \mu^*) = \min_{x \in \mathbb{R}^n} L(x, \lambda^*, \mu^*), \tag{2}$$

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, \ldots, m. \tag{3}$$

*The set of pairs $(\lambda^*, \mu^*)$ satisfying the above conditions coincide with the solutions of the dual problem (D), and with the Lagrange multipliers of the primal problem (P).*

*Proof.* Let $x^*$ be feasible, i.e., $x^* \in X$. Assume that there exists $(\lambda^*, \mu^*) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ satisfying (2) and (3). Then, for all $x \in \mathbb{R}^n$, we have

$$\begin{aligned} f(x^*) &= f(x^*) + \langle \lambda^*, g(x^*) \rangle + \langle \mu^*, h(x^*) \rangle \\ &= L(x^*, \lambda^*, \mu^*) \\ &\leq L(x, \lambda^*, \mu^*), \end{aligned} \tag{4}$$

where the first equality comes from (3) and the fact that $h(x^*) = 0$, the second one is just the definition of Lagrange function, and the last inequality follows from (2). For all $x \in X$, we also obtain

$$L(x, \lambda^*, \mu^*) = f(x) + \langle \lambda^*, g(x) \rangle + \langle \mu^*, h(x) \rangle \leq f(x), \tag{5}$$

where the inequality holds because $h(x) = 0$, $g(x) \leq 0$ and $\lambda^* \geq 0$. Therefore, the above inequality, together with (4) shows that $f(x^*) \leq f(x)$ for all $x \in X$, i.e., $x^*$ is an optimal solution for (P).

Now, assume that $x^*$ is a solution of (P). In particular, $f(x^*) > -\infty$. From [Class 5, Theorem 3] (the strong duality theorem), the dual problem (D) has a solution $(\lambda^*, \mu^*) \in D$ and there is no duality gap, i.e.,

$$f(x^*) = \omega(\lambda^*, \mu^*) = \inf_{x \in \mathbb{R}^n} L(x, \lambda^*, \mu^*) \leq L(x^*, \lambda^*, \mu^*), \tag{6}$$

where the last inequality holds from the definition of infimum. Similarly to (5), we also have

$$L(x^*, \lambda, \mu) \leq f(x^*) \quad \text{for all } (\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^p,$$

In particular, $L(x^*, \lambda, \mu) \leq f(x^*)$ holds which, together with (6) shows that (2) holds. □

Note that when $f$, $g$ and $h$ are differentiable, the condition (2) is equivalent to

$$\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*) = 0$$

because the problem $\min_x L(x, \lambda^*, \mu^*)$ is convex under the assumptions of the theorem. Therefore, the KKT conditions [Class 3, Theorem 1] are clear.

## 2. Separable problems

Assume that a vector $x$ can be written as

$$x := (x^1, x^2, \ldots, x^\ell), \quad \text{with } x^i \in \mathbb{R}^{n_i}, i = 1, \ldots, \ell.$$

Consider once again problem (P), where

$$f(x) = \sum_{i=1}^{\ell} f_i(x^i), \quad g(x) = \sum_{i=1}^{\ell} g_i(x^i), \quad h(x) = \sum_{i=1}^{\ell} h_i(x^i),$$

with $f_i \colon \mathbb{R}^{n_i} \to \mathbb{R}$, $g_i \colon \mathbb{R}^{n_i} \to \mathbb{R}^m$, $h_i \colon \mathbb{R}^{n_i} \to \mathbb{R}^p$. We say that such a problem has a *separable* structure, because each $f_i$, $g_i$ and $h_i$ does not depend on the whole vector $x$, but only on $x^i$. Even if the problem is separable, it is easy to see that this problem cannot be solved minimizing in $\mathbb{R}^{n_i}$ independently, because of the inequality constraints. However, the dual function is given as

$$
\begin{aligned}
\omega(\lambda,\mu) &= \inf_{x\in\mathbb{R}^n} \Big(f(x) + \langle \lambda, g(x)\rangle + \langle \mu, h(x)\rangle\Big) \\
&= \inf_{x\in\mathbb{R}^n} \left(\sum_{i=1}^{\ell} \Big(f_i(x^i) + \langle \lambda_i, g_i(x)\rangle + \langle \mu_i, h_i(x)\rangle\Big)\right) \\
&= \sum_{i=1}^{\ell} \left(\inf_{x^i\in\mathbb{R}^{n_i}} \Big(f_i(x^i) + \langle \lambda_i, g_i(x)\rangle + \langle \mu_i, h_i(x)\rangle\Big)\right).
\end{aligned}
$$

Defining

$$\omega_i(\lambda,\mu) := \inf_{x^i\in\mathbb{R}^{n_i}} \Big(f_i(x^i) + \langle \lambda_i, g_i(x)\rangle + \langle \mu_i, h_i(x)\rangle\Big),$$

we have $\omega(\lambda,\mu) = \sum_{i=1}^{\ell} \omega_i(\lambda,\mu)$. Thus, the objective dual function consists in $\ell$ independent problems in dimension $n_i$. When $n$ is much larger that $n_i$, $i = 1,\dots,\ell$, this last formulation can be advantageous from the computational point of view.

## 3. Robust optimization

Consider the following optimization problem:

$$
\begin{aligned}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & g(x,u) \le 0,
\end{aligned}
$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$, $g\colon \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$, $x \in \mathbb{R}^n$ is the decision variable and $u \in U \subset \mathbb{R}^p$ is a parameter. We now think in the case that $u$ varies in the set $U$, which is called *uncertainty set*. The following problem considers all the constraints for all possible values of $u$:

$$
\begin{aligned}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & g(x,u) \le 0 \quad \text{for all } u \in U.
\end{aligned}
$$

For a general set $U$, the above problem has infinite number of constraints. This means that it is difficult to solve such a problem. However, depending on the structure of $U$ and $g$, we will observe that the problem can be reformulated as an easier problem. Let us first note that the above problem is equivalent to

$$
\begin{aligned}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & \max_{u\in U} g(x,u) \le 0.
\end{aligned}
\tag{7}
$$

In this sense, we are looking for an optimal for the "worst-case scenario". Let us consider the following example:

$$g(x,u) := (a+u)^\top x, \quad U := \{u \in \mathbb{R}^m \mid A^\top u = c, u \ge 0\}.$$

Clearly, $g$ is linear with respect to $x$ and $U$ is a polyhedron. In this case, the maximization problem of the constraint in (7) is given by

$$\begin{aligned} \max_{u} \quad & (a+u)^\top x \\ \text{s.t.} \quad & A^\top u = c \\ & u \geq 0. \end{aligned}$$

Since $a^\top x$ is just a scalar in the above problem, we can rewrite it as

$$\begin{aligned} \max_{u} \quad & u^\top x \\ \text{s.t.} \quad & A^\top u = c \\ & u \geq 0. \end{aligned}$$

The dual of this problem is given by

$$\begin{aligned} \min_{\lambda} \quad & c^\top \lambda \\ \text{s.t.} \quad & A\lambda \geq x. \end{aligned}$$

Because the above problems are linear, the strong duality holds. It means that the robust problem (7) can be rewritten as

$$\begin{aligned} \min_{x,\lambda} \quad & f(x) \\ \text{s.t.} \quad & a^\top x + c^\top \lambda \leq 0 \\ & A\lambda \geq x. \end{aligned}$$

Assuming that $f$ is linear, this problem is just a linear programming problem. So, we replaced a difficult problem (7) with an easy problem by using duality tricks. Other well-known cases are when $U$ is an ellipsoid and $g$ is linear or quadratic. In such cases, we obtain second-order cone programming or semidefinite programming problems.

---

Operations Research, Advanced (Graduate School of Informatics, Kyoto University)
1st part by Ellen H. Fukuda (e-mail: ellen(at)i.kyoto-u.ac.jp, where (at) = @)