26:711:564 **OPTIMIZATION MODELS IN FINANCE**
Professor Andrzej Ruszczyński

Duality

1 The Dual Problem

Let us consider the nonlinear programming problem

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & g_i(x) \le 0, \quad i = 1, \dots, m, \\ & h_i(x) = 0, \quad i = 1, \dots, p, \\ & x \in X_0, \end{aligned} \tag{1}$$

with functions $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, p$, and with a set $X_0 \subset \mathbb{R}^n$.

Consider the Lagrangian,

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{p} \mu_i h_i(x), \tag{2}$$

as a function of *both* $x \in X_0$ and $(\lambda, \mu) \in \Lambda_0$, where

$$\Lambda_0 = \mathbb{R}^m_+ \times \mathbb{R}^p.$$

We define the *primal function* as

$$L_P(x) = \sup_{(\lambda,\mu)\in\Lambda_0} L(x, \lambda, \mu), \tag{3}$$

and the *dual function* as

$$L_D(\lambda, \mu) = \inf_{x\in X_0} L(x, \lambda, \mu). \tag{4}$$

If the supremum in (3) is $+\infty$ we set $L_P(x) = +\infty$. Similarly, if the infimum in (4) is $-\infty$ we set $L_D(\lambda, \mu) = -\infty$. Thus, we consider the primal function

L_P and the dual function L_D as functions attaining values in the extended real line $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.

The *primal problem* is to find

$$\min_{x \in X_0} L_P(x), \tag{5}$$

and the *dual problem* is to find

$$\max_{(\lambda,\mu) \in \Lambda_0} L_D(\lambda, \mu). \tag{6}$$

The duality theory investigates relations between the primal problem and the dual problem.

Let us calculate the primal function at each $x \in X_0$. If x satisfies all constraints of problem (1) then the terms in (2) that depend on λ are non-positive and the μ-terms are identically 0, and therefore $L_P(x) = f(x)$. If at least one constraint is violated, e.g. $g_j(x) > 0$, then increasing λ_j in (3) we can obtain arbitrarily large values of $\lambda_j g_j(x)$. Thus $L_P(x) = +\infty$ in this case. Summing up,

$$L_P(x) = \begin{cases} f(x) & \text{if } x \text{ is feasible for (1),} \\ +\infty & \text{otherwise.} \end{cases} \tag{7}$$

We conclude that the primal problem (5) is equivalent to the original problem (1). The dual function, however, is more difficult to calculate in an explicit form. In some cases, like linear or quadratic programming, a closed form of the dual problem can be derived.

Example 1.1 Consider the linear programming problem

$$\begin{aligned} &\min \langle c, x \rangle \\ &\text{subject to } Ax \geq b, \\ &\qquad\qquad x \geq 0, \end{aligned} \tag{8}$$

with a matrix A of dimension $m \times n$, $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. In (1) we have $f(x) = \langle c, x \rangle$, $g(x) = b - Ax$ and $X_0 = \mathbb{R}^n_+$. We formulate the Lagrangian:

$$L(x, \lambda) = \langle c, x \rangle + \langle \lambda, b - Ax \rangle = \langle c - A^T \lambda, x \rangle + \langle b, \lambda \rangle.$$

The dual function has the form

$$L_D(\lambda) = \langle b, \lambda \rangle + \inf_{x \geq 0} \langle c - A^T \lambda, x \rangle.$$

The infimum above is finite if and only if $c - A^T\lambda \geq 0$. In this case the infimum value of the product $\langle c - A^T\lambda, x\rangle$ over $x \geq 0$ is just zero. We conclude that

$$L_D(\lambda) = \begin{cases} \langle b, \lambda\rangle & \text{if } A^T\lambda \leq c, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem can, therefore be formulated as the linear programming problem

$$\begin{aligned} &\max\ \langle b, \lambda\rangle \\ &\text{subject to } A^T\lambda \leq c, \\ &\lambda \geq 0. \end{aligned} \tag{9}$$

A similar analysis can be carried out for a formulation of a linear programming problem with equality constraints.

It is instructive to consider the dual problem to the dual problem (9). We change the sign of the objective function to obtain a minimization problem and we use $u \in \mathbb{R}^n_+$ to denote the Lagrange multipliers associated with the constraints. We define the Lagrangian:

$$l(\lambda, u) = -\langle b, \lambda\rangle + \langle u, A^T\lambda - c\rangle = \langle Au - b, \lambda\rangle - \langle c, u\rangle.$$

The dual function can be calculated similarly to the previous case:

$$l_D(u) = \inf_{\lambda \geq 0} l(\lambda, u) = \begin{cases} -\langle c, u\rangle & \text{if } Au \geq b, \\ -\infty & \text{otherwise.} \end{cases}$$

We can thus write the dual problem as

$$\begin{aligned} &\max\ -\langle c, u\rangle \\ &\text{subject to } Au \geq b, \\ &u \geq 0. \end{aligned}$$

Changing the sign of the objective function again, we arrive at the primal problem (8).

If should be stressed that the fact that the dual problem to the dual problem coincides with the original problem is due to the bi-linear form of the Lagrangian.

2 Duality Relations

Assume that in problem (1) the function f is convex, the functions g_i are convex, the functions h_i are affine, and the set X_0 is convex. We also assume that the problem satisfies Slater's constraint qualification condition.

Theorem 2.1 *Assume that the problem (1) has a solution $\hat{x}$ with the corresponding vector of Lagrange multipliers $(\tilde{\lambda}, \tilde{\mu})$. Then $\tilde{x}$ is a solution of the primal problem, $(\tilde{\lambda}, \tilde{\mu})$ is a solution of the dual problem and the duality relation holds*

$$\min_{x \in X_0} L_P(x) = \max_{(\lambda,\mu) \in \Lambda_0} L_D(\lambda, \mu). \tag{10}$$

Basing on these results we can look for a solution of the primal problem by first solving the dual problem to get $(\hat{\lambda}, \hat{\mu})$ and then determining the primal solution $\hat{x}$ from the saddle point conditions.

Theorem 2.2 *Assume that the duality relation (10) holds true. If $(\hat{\lambda}, \hat{\mu}) \in \Lambda_0$ is a feasible solution of the dual problem, then every point $\hat{x} \in X_0$ such that*

(a) $L(\hat{x}, \hat{\lambda}, \hat{\mu}) = \min_{x \in X_0} L(x, \hat{\lambda}, \hat{\mu})$;

(b) all constraints of (1) are satisfied at $\hat{x}$;

(c) $\hat{\lambda}_i g_i(\hat{x}) = 0,\ i = 1, \ldots, m,$

is a global minimum of problem (1).

3 Decomposition

An important application of the duality theory is the decomposition of large scale optimization problems. Consider the nonlinear optimization problem

$$\begin{aligned} \min\ & f(x) \\ \text{subject to } & g_i(x) \le b_i, \quad i = 1, \ldots, m, \\ & h_i(x) = d_i, \quad i = 1, \ldots, p, \\ & x \in X_0, \end{aligned}$$

with $b \in \mathbb{R}^m$ and $d \in \mathbb{R}^p$. Assume that we can partition the vector x as

$$x = (x^1, \dots, x^K), \quad x_k \in \mathbb{R}^{n_k}, \quad \sum_{k=1}^{K} n_k = n,$$

in such a way that the objective and the constraint functions can be represented for all x as sums:

$$\begin{aligned} f(x) &= \sum_{k=1}^{K} f^k(x^k), \\ g_i(x) &= \sum_{k=1}^{K} g_i^k(x^k), \quad i = 1, \dots, m, \\ h_i(x) &= \sum_{k=1}^{K} h_i^k(x^k), \quad i = 1, \dots, p. \end{aligned}$$

Moreover, we assume that

$$X_0 = X_0^1 \times \dots X_0^K, \quad X_0^k \subset \mathbb{R}^{n_k}, \quad k = 1, \dots, K.$$

The Lagrangian has the form

$$\begin{aligned} L(x, \lambda, \mu) &= f(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i) + \sum_{i=1}^{p} \mu_i (h_i(x) - d_i) \\ &= \sum_{k=1}^{K} \big(f^k(x^k) + \sum_{i=1}^{m} \lambda_i g_i^k(x^k) + \sum_{i=1}^{p} \mu_i h_i^k(x^k) \big) - \langle \lambda, b \rangle - \langle \mu, d \rangle. \end{aligned}$$

Therefore the dual function can be calculated as follows

$$\begin{aligned} L_D(\lambda, \mu) &= \min_{x \in X_0} L(x, \lambda, \mu) \\ &= \sum_{k=1}^{K} \min_{x^k \in X_0^k} \Big[f^k(x^k) + \sum_{i=1}^{m} \lambda_i g_i^k(x^k) + \sum_{i=1}^{p} \mu_i h_i^k(x^k) \Big] - \langle \lambda, b \rangle - \langle \mu, d \rangle. \end{aligned}$$

It follows that the calculation of the dual function decomposes into K smaller problems, each for the subvector x^k, $k = 1, \dots, K$:

$$L_D^k(\lambda, \mu) = \min_{x^k \in X_0^k} \Big[f^k(x^k) + \sum_{i=1}^{m} \lambda_i g_i^k(x^k) + \sum_{i=1}^{p} \mu_i h_i^k(x^k) \Big].$$

In some cases these problems can be solved in a closed form, as in the examples below. In other cases, very efficient numerical methods can be employed for their solution. In any case, solving the dual problem

$$\max_{\substack{\lambda \geq 0 \\ \mu \in \mathbb{R}^p}} \Big[\sum_{k=1}^{K} L_D^k(\lambda, \mu) - \langle \lambda, b \rangle - \langle \mu, d \rangle \Big]$$

may be much easier than solving the primal problem. If the duality relation is satisfied, the above decomposition approach, together with Theorem 2.2, provides the optimal solution of the primal problem. If duality does not hold, a lower bound for the optimal value of the primal problem can be obtained (see the next section).

Example 3.1 Suppose that n power plants have to satisfy jointly some demand $b > 0$ for power. We assume that each plant j can generate power x_j between 0 and some upper bound u_j at the cost (per time unit) equal to

$$f_j(x_j) = c_j x_j + \frac{q_j}{2}(x_j)^2, \quad j = 1, \ldots, n.$$

The coefficients c_j and q_j are assumed to be positive for all j, and thus the cost functions are strictly convex. Moreover, we assume that $\sum_{j=1} u_j \geq b$, in order to be able to satisfy the demand. To satisfy it at a minimal cost, we formulate the optimization problem

$$\begin{aligned} \min \ & \sum_{j=1}^{n} f_j(x_j) \\ \text{subject to } & \sum_{j=1}^{n} x_j \geq b, \\ & 0 \leq x_j \leq u_j, \quad j = 1, \ldots, n. \end{aligned}$$

The feasible set is compact and nonempty, because $\sum_{j=1} u_j \geq b$. Therefore an optimal solution must exist. All constraint functions are affine and thus the optimal solution has to satisfy necessary conditions of optimality. Since the problem is convex, the duality relation holds true.

Denoting by λ the Lagrange multiplier associated with the demand constraint, we can write the Lagrangian as follows:

$$L(x, \lambda) = \sum_{j=1}^{n} f_j(x_j) + \lambda\Big(b - \sum_{j=1}^{n} x_j\Big).$$

Thus the dual function has the form

$$\begin{aligned}L_D(\lambda) &= b\lambda + \min_{0\le x\le u} \sum_{j=1}^{n} \big(f_j(x_j) - \lambda x_j\big)\\ &= b\lambda + \sum_{j=1}^{n} \min_{0\le x_j\le u_j} \big(f_j(x_j) - \lambda x_j\big).\end{aligned}$$

Consider the subproblem:

$$L_D^j(\lambda) = \min_{0\le x_j\le u_j} \Big(c_j x_j + \frac{q_j}{2}(x_j)^2 - \lambda x_j\Big). \tag{11}$$

If we interpret the Lagrange multiplier λ as the unit price for energy (power in a time unit) paid to the plants, this subproblem has a clear meaning: choose the production level x_j to maximize the profit of plant j.

We can minimize the quadratic function of one variable in (11) by simple calculus. The optimal solution is:

$$\hat{x}_j(\lambda) = \begin{cases} 0 & \text{if } 0 \le \lambda \le c_j,\\ (\lambda - c_j)/q_j & \text{if } c_j \le \lambda \le c_j + q_j u_j,\\ u_j & \text{if } \lambda \ge c_j + q_j u_j.\end{cases}$$

We see that the power output $\hat{x}_j(\lambda)$ of plant j is a nondecreasing function of the multiplier (price) λ. It follows from the duality theory that there exists an optimal value $\hat{\lambda}$ of the price, for which the vector $\hat{x}$ with coordinates $\hat{x}_j(\hat{\lambda})$ is an optimal solution of the problem. We must, therefore, have

$$\sum_{j=1}^{n} \hat{x}_j(\hat{\lambda}) = b.$$

The equality here is necessary because of the complementarity condition (c) of Theorem 2.2. The optimal price $\hat{\lambda}$ can be thus calculated as the price for which the power plants, driven by their profits, jointly satisfy the demand.

Each dual function can be calculated explicitly,

$$L_D^j(\lambda) = \begin{cases} 0 & \text{if } 0 \le \lambda \le c_j,\\ -(\lambda - c_j)^2/(2q_j) & \text{if } c_j \le \lambda \le c_j + q_j u_j,\\ (c_j - \lambda)u_j + q_j (u_j)^2/2 & \text{if } \lambda \ge c_j + q_j u_j.\end{cases}$$

It is non-positive for all $\lambda \geq 0$, which means that no plant has losses (we cannot force them to have losses), and the plants that produce power have positive profits. The dual problem

$$\max_{\lambda \geq 0} \left[b\lambda + \sum_{j=1}^{n} L_D^j(\lambda) \right]$$

can be interpreted as follows: maximize the value of power, $b\lambda$, minus the profits of the plants (recall that $-L_D^j$ is the profit of plant j).

If the cost functions of the plants are not convex, or if their feasible sets are disconnected, for example, $x_j = 0$ or $l_j \leq x_j \leq u_j$, the optimal solution of the dual problem provides a lower bound for the optimal value of the primal problem.

Example 3.2 There are n horses in a race. For every horse k we know the probability p_k that it wins and the amount s_k that the rest of the public is betting on it. The track keeps a certain proportion $C \in (0, 1)$ of the total amount bet and distributes the rest among the public in proportion to the amounts bet on the winning horse. We want to place bets totaling b dollars to maximize the expected net return.

Let us denote by x_k the amount bet on horse k. If horse k wins the race, we gain

$$F_k(x) = \frac{Ax_k}{x_k + s_k},$$

where $A := (1-C)(b + \sum_{i=1}^{n} s_i)$ is the total amount to be split among the winners. We can now write the corresponding optimization problem as follows:

$$\min \; -A \sum_{k=1}^{n} \frac{p_k x_k}{x_k + s_k} \tag{12}$$

$$\text{subject to} \sum_{k=1}^{n} x_k = b, \tag{13}$$

$$x \geq 0. \tag{14}$$

Note that each function

$$-\frac{x_k}{x_k + s_k} = \frac{s_k}{x_k + s_k} - 1$$

is convex in x_k. Therefore, (12)–(14) is a convex problem. Clearly its feasible set is nonempty and bounded, and hence it has an optimal solution. Since, in fact, the objective function is strictly convex, problem (12)–(14) possesses a unique optimal solution.

The constraints are affine and thus the dual of (12)–(14) has a nonempty set of optimal solutions. There is no duality gap between these problems. We will be able to write their optimal solutions explicitly.

Denoting by μ the multiplier associated with the budget constraint (13), we can write the Lagrangian:

$$\begin{aligned} L(x,\mu) &= -A\sum_{k=1}^{n}\frac{p_k x_k}{x_k+s_k} + \mu\Big(\sum_{k=1}^{n} x_k - b\Big) \\ &= \sum_{k=1}^{n}\Big(\mu x_k - A\frac{p_k x_k}{x_k+s_k}\Big) - b\mu. \end{aligned}$$

To calculate the dual function we observe that

$$\begin{aligned} L_D(\mu) &= \min_{x\geq 0}\sum_{k=1}^{n}\Big(\mu x_k - A\frac{p_k x_k}{x_k+s_k}\Big) - b\mu \\ &= \sum_{k=1}^{n}\min_{x_k\geq 0}\Big(\mu x_k - A\frac{p_k x_k}{x_k+s_k}\Big) - b\mu. \end{aligned}$$

Thus the minimization can be carried out for each k separately. Two cases are possible: if the unconstrained minimum of the expression

$$L_k(x_k,\mu) = \mu x_k - A\frac{p_k x_k}{x_k+s_k}$$

is attained at a positive x_k, then this is the constrained minimum as well. Otherwise the constrained minimum is zero. Simple calculations yield the minimum as a function of μ:

$$\hat{x}_k(\mu) = \max\left(0, \sqrt{\frac{Ap_k s_k}{\mu}} - s_k\right), \quad k = 1,\ldots,n.$$

This means that $\hat{x}_k(\mu) > 0$ if and only if

$$\sqrt{\frac{Ap_k s_k}{\mu}} > s_k,$$

which can be rewritten as follows:

$$\frac{p_k}{s_k} > \frac{\mu}{A}. \tag{15}$$

Ordering the horses (and scenarios) in such a way that

$$\frac{p_1}{s_1} \geq \frac{p_2}{s_2} \geq \cdots \geq \frac{p_n}{s_n}$$

we see that there must exist l (which depends on μ) such that

$$\hat{x}_k(\mu) = \begin{cases} \sqrt{\dfrac{A p_k s_k}{\mu}} - s_k, & k = 1, \ldots, l, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

At the optimal value $\hat{\mu}$ of the Lagrange multiplier the corresponding solution $\hat{x}(\hat{\mu})$ is feasible for (13). For every l this yields the candidate value of the multiplier:

$$\hat{\mu} = \frac{A\Big(\sum\limits_{k=1}^{l} \sqrt{p_k s_k}\Big)^2}{\Big(b + \sum\limits_{k=1}^{l} s_k\Big)^2}. \tag{17}$$

This value should guarantee that (15) is true only for $k = 1, \ldots, l$. To ensure that, we find l as the smallest integer for which

$$\sqrt{\frac{p_l}{s_l}} > \frac{\sum\limits_{k=1}^{l} \sqrt{p_k s_k}}{b + \sum\limits_{k=1}^{l} s_k} \geq \sqrt{\frac{p_{l+1}}{s_{l+1}}}.$$

Note that the left inequality holds for $l = 1$. If such an integer does not exist, we set $l = n$. Once l has been determined we can calculate the Lagrange multiplier by (17). The substitution into (16) renders the optimal bets.