3 Shadow Prices and Lagrangian Duality

3.1 Shadow Prices

A more intuitive understanding of Lagrange multipliers can be obtained by viewing (1.1) as a family of problems parameterized by $b \in \mathbb{R}^m$, the right hand side of the functional constraints. To this end, let $\phi(b) = \inf\{f(x) : h(x) = b, x \in \mathbb{R}^n\}$. It turns out that at the optimum, the Lagrange multipliers equal the partial derivatives of ϕ with respect to its parameters.

THEOREM 3.1. *Suppose that f and h are continuously differentiable on $\mathbb{R}^n$, and that there exist unique functions $x^* : \mathbb{R}^m \to \mathbb{R}^n$ and $\lambda^* : \mathbb{R}^m \to \mathbb{R}^m$ such that for each $b \in \mathbb{R}^m$, $h(x^*(b)) = b$, $\lambda^*(b) \leqslant 0$ and $f(x^*(b)) = \phi(b) = \inf\{f(x) - \lambda^*(b)^T(h(x) - b) : x \in \mathbb{R}^n\}$. If x^* and λ^* are continuously differentiable, then*

$$\frac{\partial \phi}{\partial b_i}(b) = \lambda_i^*(b).$$

Proof. We have that

$$\begin{aligned}
\phi(b) &= f(x^*(b)) - \lambda^*(b)^T(h(x^*(b)) - b) \\
&= f(x^*(b)) - \lambda^*(b)^T h(x^*(b)) + \lambda^*(b)^T b.
\end{aligned}$$

Taking partial derivatives of each term,

$$\frac{\partial f(x^*(b))}{\partial b_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x^*(b)) \frac{\partial x_j^*}{\partial b_i}(b),$$

$$\begin{aligned}
\frac{\partial \lambda^*(b)^T h(x^*(b))}{\partial b_i} &= \lambda^*(b)^T \frac{\partial h(x^*(b))}{\partial b_i} + h(x^*(b)) \frac{\partial \lambda^*(b)^T}{\partial b_i} \\
&= \left(\sum_{j=1}^{n} \left(\lambda^*(b)^T \frac{\partial h}{\partial x_j}(x^*(b)) \right) \frac{\partial x_j^*}{\partial b_i}(b) \right) + h(x^*(b)) \frac{\partial \lambda^*(b)^T}{\partial b_i},
\end{aligned}$$

$$\frac{\partial \lambda^*(b)^T b}{\partial b_i} = \lambda^*(b)^T \frac{\partial b}{\partial b_i} + b \frac{\lambda^*(b)^T}{\partial b_i}.$$

By summing and re-arranging,

$$\begin{aligned}
\frac{\partial \phi(b)}{\partial b_i} = \sum_{j=1}^{n} \left(\frac{\partial f}{\partial x_j}(x^*(b)) - \lambda^*(b)^T \frac{\partial h}{\partial x_j}(x^*(b)) \right) \frac{\partial x_j^*}{\partial b_i}(b) \\
- (h(x^*(b)) - b) \frac{\partial \lambda^*(b)^T}{\partial b_i} + \lambda^*(b)^T \frac{\partial b}{\partial b_i}.
\end{aligned}$$

The first term on the right-hand side is zero, because $x^*(b)$ minimizes $L(x, \lambda^*(b))$ and thus

$$\frac{\partial L(x^*(b), \lambda^*(b))}{\partial x_j} = \frac{\partial f}{\partial x_j}(x^*(b)) - \left(\lambda^*(b)^T \frac{\partial h}{\partial x_j}(x^*(b))\right) = 0$$

for $j = 1, \dots, n$. The second term is zero as well, because $x^*(b)$ is feasible and thus $(h(x^*(b)) - b)_k = 0$ for $k = 1, \dots, m$, and the claim follows. □

It should be noted that the result also holds when the functional constraints are inequalities: if the ith constraint does not not hold with equality, then $\lambda_i^* = 0$ by complementary slackness, and therefore also $\partial\lambda_i^*/\partial b_i = 0$.

The Lagrange multipliers are also known as *shadow prices*, due to an economic interpretation of the problem to

$$\begin{aligned} \text{maximize} \quad & f(x) \\ \text{subject to} \quad & h(x) \leqslant b \\ & x \in X. \end{aligned}$$

Consider a firm that produces n different goods from m different raw materials. Vector $b \in \mathbb{R}^m$ describes the amount of each raw material available to the firm, vector $x \in \mathbb{R}^n$ the quantity produced of each good. Functions $h : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^n \to \mathbb{R}$ finally describe the amounts of raw material required to produce, and the profit derived from producing, particular quantities of the goods. The goal of the above problem thus is to maximize the profit of the firm for given amounts of raw materials available to it. The *shadow price* of raw material i then is the price the firm would be willing to pay per additional unit of this raw material, which of course should be equal to the additional profit derived from it, i.e., to $\frac{\partial \phi}{\partial b_i}(b)$.

3.2 Lagrangian Duality

Another useful concept that arises from Lagrange multipliers is that of a dual problem. Again consider the optimization problem (1.1) and the Lagrangian (2.1), and define the (Lagrange) *dual function* $g : \mathbb{R}^m \to \mathbb{R}$ as the minimum value of the Lagrangian over X, i.e.,

$$g(\lambda) = \inf_{x \in X} L(x, \lambda). \tag{3.1}$$

As before, let Y be the set vectors of Lagrange multipliers for which the Lagrangian has a finite minimum, i.e., $Y = \{\lambda \in \mathbb{R}^m : \inf_{x \in X} L(x, \lambda) > -\infty\}$.

It is easy to see that the maximum value of the dual function provides a lower bound on the minimum value of the original objective function. This property is known as weak duality.

THEOREM 3.2 (Weak duality). *If $x \in X(b)$ and $\lambda \in Y$, then $g(\lambda) \leqslant f(x)$, and in particular,*

$$\sup_{\lambda \in Y} g(\lambda) \leqslant \inf_{x \in X(b)} f(x). \tag{3.2}$$

Proof. Let $x \in X(b)$ and $\lambda \in Y$. Then,

$$\begin{aligned} g(\lambda) &= \inf_{x' \in X} L(x', \lambda) \\ &\leqslant L(x, \lambda) \\ &= f(x) - \lambda^T(h(x) - b) \\ &= f(x). \end{aligned}$$

□

Equality on the first and third line holds by definition of g and L, the inequality on the second line because $x \in X$. The last equality holds because $x \in X(b)$ and therefore $h(x) - b = 0$.

In light of this result, it is interesting to choose λ in order to make this lower bound as large as possible, i.e., to

$$\begin{aligned} &\text{maximize} \quad g(\lambda) \\ &\text{subject to} \quad \lambda \in Y. \end{aligned}$$

This problem is known as the *dual problem*, and (1.1) is in this context referred to as the *primal problem*. If (3.2) holds with equality, i.e., if there exists $\lambda \in Y$ such that $g(\lambda) = \inf_{x \in X(b)} f(x)$, the problem is said to satisfy *strong duality*. The cases where strong duality holds are those that can be solved using the method of Lagrange multipliers.

EXAMPLE 3.3. Again consider the minimization problem of Example 2.2, and recall that $Y = \{\lambda \in \mathbb{R}^2 : \lambda_1 = -2, \lambda_2 < 0\}$ and that for each $\lambda \in Y$ the minimum occurred at $x^*(\lambda) = (3/(2\lambda_2), 1/(2\lambda_2), x_3)$. Thus,

$$g(\lambda) = \inf_{x \in X} L(x, \lambda) = L(x^*(\lambda), \lambda) = \frac{10}{4\lambda_2} + 4\lambda_2 - 10,$$

so the dual problem is to

$$\text{maximize } \frac{10}{4\lambda_2} + 4\lambda_2 - 10 \text{ subject to } \lambda_2 < 0.$$

It should not come as a surprise that the maximum is attained for $\lambda_2 = -\sqrt{5/8}$, and that the primal and dual have the same optimal value, namely $-2(\sqrt{10}+5)$. Note that it is not actually necessary to solve the dual to see that $\lambda_2 = -\sqrt{5/8}$ is an optimizer, it suffices that the value of the dual function at this point equals the value of the objective function of the primal at some point in the feasible set of the primal.

There are several reasons why the dual is interesting. Any feasible solution of the dual provides a succinct certificate that the optimal solution of the primal is bounded by a certain value. In particular, a pair of solutions of the primal and dual that yield the same value must be optimal. If strong duality holds, the optimal value of the primal can be determined by solving the dual, which in some cases may be easier than solving the primal. In a later lecture we will express two quantities as the optimal solutions of a pair of a primal and a dual that satisfy strong duality, thereby showing that the two quantities are equal.