Chapter 12
Sensitivity Analysis and Envelope Theorems

12.1 The Meaning of the Multipliers

The Lagrange multipliers will be shown to be the rate of change of the optimal value of the objective function per unit change of the RHS's of the constraints. They are called the ***economic values***, ***shadow prices*** or ***opportunity costs*** of the RHS's.

We will discuss the Lagrange multiplier for the one equality constraint case and then, with the insights gained, generalize to the other cases.

12.2 The Meaning of Lagrange Multiplier: One Equality Constraint Case

Theorem (**Simon and Blume** [1994], Theorem 19.1, page 449) Consider the optimization problem

$$\begin{gathered}\max(\min) f(\mathbf{x}) \\ \text{st.}\, g(\mathbf{x}) = a,\end{gathered}$$

where *f* and *g* are functions in $\boldsymbol{C}^1$. For any fixed a, let $\mathbf{x}(a)$ be a local optimal solution to the optimization problem, with corresponding multiplier $\lambda(a)$. So $\mathbf{x}: \mathbb{R} \to \mathbb{R}^n$ and $\lambda: \mathbb{R} \to \mathbb{R}$ are functions and suppose that they are in $\boldsymbol{C}^1$. For a given value of a,

$$\lambda(a) = \frac{df(\mathbf{x}(a))}{da}$$

Proof The composite function $M(a) = f(\mathbf{x}(a))$ is in $\boldsymbol{C}^1$ because $f(\mathbf{x})$ and $\mathbf{x}(a)$ are. So its derivative exists. Now we will show that the derivative of *F* with respect to *a* is equal to the multiplier. Write the Lagrange function, with *a* being the parameter,

$$\mathcal{L}(\mathbf{x}, \lambda; a) = f(\mathbf{x}) - \lambda(g(\mathbf{x}) - a).$$

The first-order necessary condition is

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}(a), \lambda(a); a) = \nabla f(\mathbf{x}(a)) - \lambda(a)\nabla g(\mathbf{x}(a)) = \mathbf{0}.$$

Since $g\big(\mathbf{x}(a)\big) = a$, for any a, the derivative of g with respect to a by chain rule is

$$\frac{d}{da} g\big(\mathbf{x}(a)\big) = \nabla g\big(\mathbf{x}(a)\big)^{\mathrm{T}}\mathbf{x}'(a) = 1$$

Then take the derivative of $f(\mathbf{x}(a))$ with respect to a by chain rule, we have

$$\frac{d}{da}M(a) = \frac{d}{da}f(\mathbf{x}(a)) = \nabla f(\mathbf{x}(a))^{\mathrm{T}}\mathbf{x}'(a)$$
$$= \lambda(a)\nabla g(\mathbf{x}(a))^{\mathrm{T}}\mathbf{x}'(a)$$
$$= \lambda(a). \blacksquare$$

This theorem is equally applicable in maximization and minimization problems.

HW. What is the relationship between the Lagrange multiplier λ_a of the problem

$$\max f(\mathbf{x})$$
$$\text{st.}\, g(\mathbf{x}) = a,$$

and the Lagrange multiplier λ_a of the problem

$$\min g(\mathbf{x})$$
$$\text{st.}\, f(\mathbf{x}) = b,$$

if we assume that $f(\mathbf{x}(a) = b$.

Example In the problem of maximizing the utility function $u(\mathbf{x})$ subject to the budget constraint $\mathbf{p}^{\mathrm{T}}\mathbf{x} = B$, the Lagrange multiplier $\lambda(B)$ is the marginal utility the consumer receives per one additional baht of money. If he has ΔB bahts more, he will have approximately $\Delta u \approx \lambda(B)\Delta B$ additional utility. (Note this meaning and refer back to the inter-temporal consumption example in Section 9.3.1.)

HW In the problem of minimizing the expenditure $\mathbf{p}^{\mathrm{T}}\mathbf{x} = B$ subject to the level of utility $u(\mathbf{x}) = u_0$, interpret the economic meaning of the Lagrange multiplier $\mu(u_0)$. How are $\mu(u_0)$ and $\lambda(B)$ of the previous example related if the maximal utility attained in the previous example $u(\mathbf{x}(B)) = u_0$?

Example Simon and Blume [1994], Example 19.1. The problem of maximizing $x_1^2x_2$ subject to the constraint $2x_1^2 + x_2^2 = 3$ has the Lagrange multiplier $\lambda(3) = 0.5$. If the RHS is changed to 3.3, the optimal objective function will change by approximately $0.5 \times 0.3 = 0.15$.

12.3 The Meaning Lagrange Multiplier: Several Equality Constraint Case

Theorem (**Simon and Blume** [1994], Theorem 19.2, page 450) Consider the optimization problem with k equality constraints,

$$\max\,(\min) f(\mathbf{x})$$
$$\text{st.}\, \mathbf{g}(\mathbf{x}) = \mathbf{a},$$

where f and $\mathbf{g}$ are functions in $\boldsymbol{C}^1$. For any fixed vector $\mathbf{a}$, let $\mathbf{x}(\mathbf{a})$ be a local optimal solution to the optimization problem, with corresponding multiplier $\boldsymbol{\lambda}(\mathbf{a})$. So $\mathbf{x}: \mathbb{R} \to \mathbb{R}^n$ and $\boldsymbol{\lambda}: \mathbb{R}^k \to \mathbb{R}^k$ are functions and suppose that they are in $\boldsymbol{C}^1$. For a given value of $\mathbf{a}$,

$$\boldsymbol{\lambda}(\mathbf{a}) = \nabla_{\mathbf{a}} f\big(\mathbf{x}(\mathbf{a})\big).$$

Proof The composite function $F(\mathbf{a}) = f(\mathbf{x}(\mathbf{a}))$ is in $\boldsymbol{C}^1$ because f and $\mathbf{x}$ are and thus its gradient exists. We will show that it is the vector of multipliers as given in the theorem. Write the Lagrange function, with $\mathbf{a}$ being the parameter,

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}; \mathbf{a}) = f(\mathbf{x}) - \boldsymbol{\lambda}^{\mathrm{T}}(\mathbf{g}(\mathbf{x}) - \mathbf{a}).$$

The first-order necessary conditions that

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}(\mathbf{a}), \boldsymbol{\lambda}(\mathbf{a}); \mathbf{a}) = \nabla_{\mathbf{x}} f\big(\mathbf{x}(\mathbf{a})\big) - \nabla_{\mathbf{x}}\mathbf{g}\big(\mathbf{x}(\mathbf{a})\big)^{\mathrm{T}}\boldsymbol{\lambda}(\mathbf{a}) = \mathbf{0}$$

Since $\mathbf{g}\big(\mathbf{x}(\mathbf{a})\big) = \mathbf{a}$, for any $\mathbf{a}$, the derivative of $\mathbf{g}$ with respect to $\mathbf{a}$ by chain rule

$$\nabla_{\mathbf{a}}\mathbf{g}\big(\mathbf{x}(\mathbf{a})\big) = \nabla_{\mathbf{x}}\mathbf{g}\big(\mathbf{x}(\mathbf{a})\big)\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) = \mathbf{I},$$

where $\mathbf{I}$ is an identity matrix of order k. Then taking the derivative of $f(\mathbf{x}(\mathbf{a}))$ with respect to $\mathbf{a}$ by chain rule, we have

$$\begin{aligned}\nabla_{\mathbf{a}} f\big(\mathbf{x}(\mathbf{a})\big)^{\mathrm{T}} &= \nabla_{\mathbf{x}} f\big(\mathbf{x}(\mathbf{a})\big)^{\mathrm{T}}\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) \\ &= \boldsymbol{\lambda}(\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{x}}\mathbf{g}\big(\mathbf{x}(\mathbf{a})\big)\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) \\ &= \boldsymbol{\lambda}(\mathbf{a})^{\mathrm{T}}.\end{aligned}$$

❑

12.4 The Meaning of Lagrange Multiplier: Inequality Constraint Case

Theorem (**Simon and Blume** [1994], Theorem 19.3, page 451) Consider the maximization problem with m inequality constraints,

$$\max f(\mathbf{x})$$
$$\mathbf{h}(\mathbf{x}) \leq \mathbf{a},$$

where f and $\mathbf{h}$ are functions in $\boldsymbol{C}^1$. For any fixed vector $\mathbf{a}$, let $\mathbf{x}(\mathbf{a})$ be a local optimal solution to the optimization problem, with corresponding multiplier $\boldsymbol{\mu}(\mathbf{a})$. So $\mathbf{x}: \mathbb{R}^m \to \mathbb{R}^n$ and $\boldsymbol{\mu}: \mathbb{R}^m \to \mathbb{R}^m$ are functions and suppose that they are in $\boldsymbol{C}^1$. For a given value of $\mathbf{a}$,

$$\boldsymbol{\mu}(\mathbf{a}) = \nabla_{\mathbf{a}} f\big(\mathbf{x}(\mathbf{a})\big).$$

Proof (Sketch) Suppose that the first b constraints are binding, then we can apply the previous theorem to obtain

$$\mu_i(\mathbf{a}) = \frac{\partial f(\mathbf{x}(\mathbf{a}))}{\partial a_i}, i = 1,2,\dots,r.$$

We will show that the multipliers for the unbinding constraints are zeros. For the constraints that are not binding at $\mathbf{x}(\mathbf{a})$, $h^i\big(\mathbf{x}(\mathbf{a})\big) < a_i$, $i = r+1, r+2, \dots, m$. We can thus reduce the values of unbinding constraints $i = r+1, r+2, \dots, m$, by some sufficiently small $\varepsilon > 0$ and still keep the solution $\mathbf{x}(\mathbf{a})$ feasible. With this new value of RHS $\mathbf{a}'$, the new feasible set is a subset of the original one. Thus the solution $\mathbf{x}(\mathbf{a})$ is also a local maximum of the new feasible set. That is,

$$f\big(\mathbf{x}(\mathbf{a})\big) = f\big(\mathbf{x}(\mathbf{a}')\big).$$

This means for some sufficiently small change in $a_i, i = r+1, r+2, \dots, m$, there is no change in the maximal objective function value. So $\frac{\partial f(\mathbf{x}(\mathbf{a}))}{\partial a_i} = 0, i = r+1, r+2, \dots, m.$❑

HW Write the corresponding result of the previous theorem for the minimization problem.

HW Write the theorem for the mixed-constraint optimization.

Example Simon and Blume [1994], Example 19.2. The problem of maximizing xyz subject to the constraints, $x + y + z \leq 1$, $x \geq 0, y \geq 0$, and $z \geq 0$, has the optimal solution $(x^*, y^*, z^*) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Only the first constraint is binding and its Lagrange multiplier is $\frac{1}{9}$, while all others are zero. If the RHS of the first constrant is increased by 0.1, the optimal objective function value will change approximately by $0.1 \times \frac{1}{9}$.

Example At the optimal solution of the Linear Programming problem,

$$\begin{aligned} &\max f(\mathbf{x}) = \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ &\text{st.}\, \mathbf{h}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b} \le \mathbf{0} \\ &\qquad\quad \mathbf{x} \ge \mathbf{0}, \end{aligned}$$

the value of the Lagrange multiplier vector $\boldsymbol{\mu}(\mathbf{b})$ is the rate of the change of the optimal objective function value as the RHS $\mathbf{b}$ change.

HW In the dual problem,

$$\begin{aligned} &\min \mathbf{b}^{\mathrm{T}}\boldsymbol{\mu} \\ &\text{st.}\, \mathbf{A}^{\mathrm{T}}\boldsymbol{\mu} \ge \mathbf{c} \\ &\qquad\ \boldsymbol{\mu} \ge \mathbf{0}, \end{aligned}$$

let $\mathbf{x}$ be the vector of the Lagrange multipliers. Interpret the critical value of the Lagrange multiplier $\mathbf{x}^*$.

12.5 Envelope Theorems

We can generalize the analysis of the change in the optimal value of the objective function when the RHS changes. The Envelope Theorem will analyze the change in the optimal objective function value when a parameter of the problem changes. For example, in the unconstrained maximization of $f(x_1, x_2) = 3x_1^a x_2 - x_1 x_2^3$, what is the rate of change of the optimal objective function value when the parameter a changes?

12.6 Envelope Theorem: Unconstrained Case

Theorem (**Simon and Blume** [1994], Theorem 19.4) Let $f(\mathbf{x}; a)$ be a $\boldsymbol{C}^1$ function of $\mathbf{x} \in \mathbb{R}^n$ and scalar $a \in \mathbb{R}$. For each choice of a, consider the unconstrained optimization $\max(\min) f(\mathbf{x}; a)$. Let $\mathbf{x}(a)$ be a local optimal solution for each specification of a, and so $\mathbf{x}(a)$ is a function of a. Suppose that $\mathbf{x}(a)$ is a $\boldsymbol{C}^1$ function and write $M(a) = f(\mathbf{x}(a); a)$, which is called ***the maximal (minimal) value function***. Then,

$$\frac{dM(a)}{da} = \left.\frac{\partial f(\mathbf{x}; a)}{\partial a}\right|_{\mathbf{x}(a)},$$

where the partial derivative on the right hand side is evaluated at the point $\mathbf{x}(a)$.

Proof By the definition that $M(a) = f(\mathbf{x}(a); a)$, the derivate of $M(a)$ is equal to the total derivative of $f(\mathbf{x}(a); a)$, and by chain rule,

$$\begin{aligned}\frac{dM(a)}{da} &= \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}(a); a)}{\partial x_j}\frac{dx_j}{da} + \left.\frac{\partial f(\mathbf{x}; a)}{\partial a}\right|_{\mathbf{x}(a)} \\ &= \nabla_{\mathbf{x}} f(\mathbf{x}(a); a)^{\mathrm{T}}\mathbf{x}'(\mathbf{a}) + \left.\frac{\partial f(\mathbf{x}; a)}{\partial a}\right|_{\mathbf{x}(a)} \\ &= \left.\frac{\partial f(\mathbf{x}; a)}{\partial a}\right|_{\mathbf{x}(a)}.\end{aligned}$$

The last equation holds because when $\mathbf{x}(a)$ is a local optimal solution the gradient of f with respect to $\mathbf{x}$, $\nabla_{\mathbf{x}} f(\mathbf{x}(a); a)$ is necessarily zero by the first-order necessary condition.❑

Example Maximize $f(x; a) = -x^2 + 4ax + 4a^2$. The optimal solution is $x(a) = 2a$. The maximal value function is $M(a) = 8a^2$, and $M'(a) = 16a$.

Using the Envelope Theorem we also get

$$\begin{aligned}\left.\frac{\partial f(x; a)}{\partial a}\right|_{x(a)} &= 4x + 8a\Big|_{x(a)} \\ &= 4x(a) + 8a = 4 \times 2a + 8a = 16a.\end{aligned}$$

Example *Defectives in Production*. If s is the fraction of total output that is defective, the profit function becomes

$$\pi(L, K; s) = (1 - s)p \cdot f(L, K) - (wL + rK).$$

Then, the maximal value function is

$$\begin{aligned}M(s) &= \pi(L(s), K(s); s) \\ &= (1 - s)p \cdot f\big(L(s), K(s)\big) - \big(wL(s) + rK(s)\big),\end{aligned}$$

and by the Envelope Theorem,

$$\begin{aligned}\frac{dM(s)}{ds} &= \left.\frac{\partial \pi(L, K; s)}{\partial s}\right|_{L(s), K(s)} \\ &= -pf(L, K)\Big|_{L(s), K(s)} \\ &= -pf\big(L(s), K(s)\big),\end{aligned}$$

which is just the total revenue at the optimal production point. If the defective fraction s is changed by 0.01, i.e. 1%, the profit will decrease by approximately,

$$\begin{aligned}\Delta M(s) &\approx \frac{dM(s)}{ds}\cdot\Delta s\\ &= -0.01\frac{dM(s)}{ds}\\ &= -0.01pf\big(L(s),K(s)\big),\end{aligned}$$

which is 1% of the total revenue.

When the function f depends on r parameters, we can find the partial derivative of the value function with respect to each parameter.

Corollary Consider the unconstrained optimization $\max\min f(\mathbf{x};\mathbf{a})$, $\boldsymbol{x}\in\mathbb{R}^n$, *for* each choice of vector of parameters $\mathbf{a}\in\mathbb{R}^r$. Assume that $\mathbf{x}(\mathbf{a})$ be a local optimal solution that is a $\boldsymbol{C}^1$ function of $\mathbf{a}$. Write the value function $M(\mathbf{a}) = f(\mathbf{x}(\mathbf{a});\mathbf{a})$. Then,

$$\nabla_{\mathbf{a}}M(\mathbf{a}) = \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})\Big|_{\mathbf{x}(\mathbf{a})}.$$

Proof Using the chain rule,

$$\begin{aligned}\nabla_{\mathbf{a}}M(\mathbf{a})^{\mathrm{T}} &= \nabla_{\mathbf{x}}f(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) + \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a})}.\\ &= \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a})},\end{aligned}$$

where $\nabla_{\mathbf{x}}f(\mathbf{x}(\mathbf{a});\mathbf{a}) = \mathbf{0}$ by the first-order necessary condition.❑

Example Maximize the profit function $\pi(L,K) = pf(L,K) - (wL + rK)$. Taking (w,r) as parameters, by the Envelope Theorem we have

$$\begin{aligned}\frac{\partial M(w,r)}{\partial r} &= \frac{\partial \pi(L,K;w,r)}{\partial r}\Big|_{L(w,r),K(w,r)}\\ &= -K\Big|_{L(w,r),K(w,r)}\\ &= -K(w,r).\end{aligned}$$

That is, when the price of capital r changes the total profit will change equal to the level of capital used $K(w,r)$ per unit change in r.

12.7 Envelope Theorem: Equality Constraints Case

Theorem (**Simon and Blume** [1994], Theorem 19.5, and **Jehle** [1991], Theorem 2.4.1) Consider the constrained optimization,

$$\max\min f(\mathbf{x};\mathbf{a})$$
$$\text{st. } \mathbf{g}(\mathbf{x};\mathbf{a}) = \mathbf{0},$$

where $\mathbf{a} \in \mathbb{R}^r$ is a vector of parameters. Let $\mathbf{x}(\mathbf{a})$ be a local optimal solution with multiplier $\boldsymbol{\lambda}(\mathbf{a})$, for each choice of $\mathbf{a}$. Suppose that $\mathbf{x}(\mathbf{a})$ is a $\boldsymbol{C}^1$ function. Write the maximal value function $M(\mathbf{a}) = f(\mathbf{x}(\mathbf{a});\mathbf{a})$. Then,

$$\nabla_{\mathbf{a}} M(\mathbf{a}) = \nabla_{\mathbf{a}} \mathcal{L}(\mathbf{x},\boldsymbol{\lambda};\mathbf{a}) \Big|_{\mathbf{x}(\mathbf{a}),\boldsymbol{\lambda}(\mathbf{a})},$$

where $\mathcal{L}$ is the Lagrange function,

$$\mathcal{L}(\mathbf{x},\boldsymbol{\lambda};\mathbf{a}) = f(\mathbf{x};\mathbf{a}) - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{g}(\mathbf{x};\mathbf{a})$$

Proof By chain rule, as in the previous corollary,

$$\nabla_{\mathbf{a}} M(\mathbf{a})^{\mathrm{T}} = \nabla_{\mathbf{x}} f(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}} \nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) + \nabla_{\mathbf{a}} f(\mathbf{x};\mathbf{a})^{\mathrm{T}} \Big|_{\mathbf{x}(\mathbf{a})} \quad (1)$$

Since $\mathbf{x}(\mathbf{a})$ is a local optimal solution, from the first-order necessary condition there exists $\boldsymbol{\lambda}(\mathbf{a})$ that

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}(\mathbf{a}),\boldsymbol{\lambda}(\mathbf{a});\mathbf{a}) = \nabla_{\mathbf{x}} f(\mathbf{x}(\mathbf{a});\mathbf{a}) - \nabla_{\mathbf{x}}\mathbf{g}(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\boldsymbol{\lambda}(\mathbf{a}) = \mathbf{0},$$

we have,

$$\nabla_{\mathbf{x}} f(\mathbf{x}(\mathbf{a});\mathbf{a}) = \nabla_{\mathbf{x}}\mathbf{g}(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\boldsymbol{\lambda}(\mathbf{a}). \quad (2)$$

Since $\mathbf{g}(\mathbf{x};\mathbf{a}) = \mathbf{0}$ for any feasible point $\mathbf{x}$ and $\mathbf{a}$,

$$\nabla_{\mathbf{a}}\mathbf{g}(\mathbf{x}(\mathbf{a});\mathbf{a}) = \nabla_{\mathbf{x}}\mathbf{g}(\mathbf{x}(\mathbf{a});\mathbf{a})\nabla\mathbf{x}(\mathbf{a}) + \nabla_{\mathbf{a}}\mathbf{g}(\mathbf{x};\mathbf{a}) \Big|_{\mathbf{x}(\mathbf{a})} = \mathbf{0}$$

and so,

$$\nabla_{\mathbf{x}}\mathbf{g}(\mathbf{x}(\mathbf{a});\mathbf{a})\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) = -\nabla_{\mathbf{a}}\mathbf{g}(\mathbf{x};\mathbf{a}) \Big|_{\mathbf{x}(\mathbf{a})} \quad (3)$$

From (1), (2) and (3),

$$\begin{aligned}\nabla_{\mathbf{a}} M(\mathbf{a})^{\mathrm{T}} &= \boldsymbol{\lambda}(\mathbf{a})^{\mathrm{T}} \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}} \nabla_{\mathbf{a}} \mathbf{x}(\mathbf{a}) + \nabla_{\mathbf{a}} f(\mathbf{x};\mathbf{a})^{\mathrm{T}} \Big|_{\mathbf{x}(\mathbf{a})} \\ &= \boldsymbol{\lambda}(\mathbf{a})^{\mathrm{T}} \left(-\nabla_{\mathbf{a}} \mathbf{g}(\mathbf{x};\mathbf{a}) \Big|_{\mathbf{x}(\mathbf{a})} \right) + \nabla_{\mathbf{a}} f(\mathbf{x};\mathbf{a})^{\mathrm{T}} \Big|_{\mathbf{x}(\mathbf{a})} \\ &= \nabla_{\mathbf{a}} f(\mathbf{x};\mathbf{a})^{\mathrm{T}} \Big|_{\mathbf{x}(\mathbf{a})} - \boldsymbol{\lambda}(\mathbf{a})^{\mathrm{T}} \left(\nabla_{\mathbf{a}} \mathbf{g}(\mathbf{x};\mathbf{a}) \Big|_{\mathbf{x}(\mathbf{a})} \right) \\ &= \nabla_{\mathbf{a}} \mathcal{L}(\mathbf{x},\boldsymbol{\lambda};\mathbf{a})^{\mathrm{T}} \Big|_{\mathbf{x}(\mathbf{a})}.\end{aligned}$$

Example See **Simon and Blume** [1994], Example 19.6 and **Jehle** [1991], Example 2.4.1.

HW Show that the theorem on the meaning of the multipliers is just a special case of this Envelope Theorem in the equality constraint case.

12.8 Applications: Shephard's Lemma, Roy's Identity and Demand Function

Theorem (**Shephard's Lemma**) Consider the cost minimization , subject to the output level $f(\mathbf{x}) = y$, f is in $\boldsymbol{C}^1$. This problem is parameterized by $(\mathbf{w}, y)$. Let $\mathbf{x}(\mathbf{w}, y)$ be the optimal solution and assume that it is in $\boldsymbol{C}^1$. The value function $C(\mathbf{w}, y) = \mathbf{w}^{\mathrm{T}}\mathbf{x}(\mathbf{w}, y)$ is just the cost function as it is the minimum cost the firm has to pay for a given output level y at given input prices $\mathbf{w}$. Then,

$$\nabla_{\mathbf{w}} C(\mathbf{w}, y) = \mathbf{x}(\mathbf{w}, y),$$

or equivalently,

$$\frac{\partial C(\mathbf{w}, y)}{\partial w_j} = x_j(\mathbf{w}, y), j = 1,2, \ldots, n,$$

which is the demand function for input j as a function of input prices and output level.

Proof The Lagrange function is given by

$$\mathcal{L}(\mathbf{x}, \lambda; \mathbf{w}, y) = \mathbf{w}^{\mathrm{T}}\mathbf{x} - \lambda(f(\mathbf{x}) - y)$$

By the Envelope Theorem,

$$\nabla_{\mathbf{w}} C(\mathbf{w}, y) = \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{x}, \lambda; \mathbf{w}, y) \Big|_{\mathbf{x}(\mathbf{w}, y), \lambda(\mathbf{w}, y)}$$

$$= \nabla_{\mathbf{w}}(\mathbf{w}^{\mathrm{T}}\mathbf{x} - \lambda(f(\mathbf{x}) - y)) \Big|_{\mathbf{x}(\mathbf{w},y),\lambda(\mathbf{w},y)}$$
$$= \mathbf{x}\Big|_{\mathbf{x}(\mathbf{w},y),\lambda(\mathbf{w},y)}$$
$$= \mathbf{x}(\mathbf{w},y).$$

❑

HW What is the partial derivative of the cost function with respect to output?

Theorem (**Indirect Utility Function and Demand**) For the consumer's problem that a utility function is maximized subject to the budget constraint , let $\mathbf{x}(\mathbf{p},b)$ be the optimal solution for given prices $\mathbf{p}$ and income b. The value function here is $v(\mathbf{p},b) = u(\mathbf{x}(\mathbf{p},b);\mathbf{p},b)$, which is called the ***indirect utility function***. Suppose u and $\mathbf{x}(\mathbf{p},b)$ are in $\boldsymbol{C}^1$, then

$$x_j(\mathbf{p},b) = -\frac{\frac{\partial v(\mathbf{p},b)}{\partial p_j}}{\frac{\partial v(\mathbf{p},b)}{\partial b}}, j = 1,2,\dots,n,$$

and this is called the ***Roy's Identity***.

Proof Write the Lagrange function

$$\mathcal{L}(\mathbf{x},\lambda;\mathbf{p},b) = u(\mathbf{x}) - \lambda(\mathbf{p}^{\mathrm{T}}\mathbf{x} - b)$$

and apply the Envelope Theorem, we have

$$\nabla_{\mathbf{p}} v(\mathbf{p},b) = \nabla_{\mathbf{p}}\mathcal{L}(\mathbf{x},\lambda;\mathbf{p},b)\Big|_{\mathbf{x}(\mathbf{p},b),\lambda(\mathbf{p},b)}$$
$$= \lambda\mathbf{x}\Big|_{\mathbf{x}(\mathbf{p},b),\lambda(\mathbf{p},b)}$$
$$= \lambda(\mathbf{p},b)\mathbf{x}(\mathbf{p},b)\Big|_{\mathbf{x}(\mathbf{p},b),\lambda(\mathbf{p},b)},$$

and

$$\frac{\partial v(\mathbf{p},b)}{\partial b} = \frac{\partial \mathcal{L}(\mathbf{x},\lambda;\mathbf{p},b)}{\partial b}\Big|_{\mathbf{x}(\mathbf{p},b),\lambda(\mathbf{p},b)}$$
$$= \lambda(\mathbf{p},b).$$

❑

HW Derive the ***Hicksian demand function*** $\mathbf{x}(\mathbf{p},a)$, from the minimization of the expenditure $\mathbf{p}^{\mathrm{T}}\mathbf{x}$, subject to the utility level $u(\mathbf{x}) = a$.

12.9 Envelope Theorem: Inequality Constraints Case

We can show that the Envelope Theorem still remains the same under inequality constraints.

Theorem Consider the constrained optimization,

$$\begin{aligned}&\max f(\mathbf{x};\mathbf{a})\\&\text{st.}\,\mathbf{h}(\mathbf{x};\mathbf{a})\le \mathbf{0},\end{aligned}$$

where $\mathbf{a}\in\mathbb{R}^r$ is a vector of parameters. Let $\mathbf{x}(\mathbf{a})$ be a local optimal solution with multiplier $\boldsymbol{\mu}(\mathbf{a})$, for each choice of $\mathbf{a}$. Suppose that $\mathbf{x}(\mathbf{a})$ is a $\boldsymbol{C}^1$ function. Write the value function $M(\mathbf{a})=f(\mathbf{x}(\mathbf{a});\mathbf{a})$. Then,

$$\nabla_{\mathbf{a}}M(\mathbf{a})=\nabla_{\mathbf{a}}\mathcal{L}(\mathbf{x},\boldsymbol{\mu};\mathbf{a})\Big|_{\mathbf{x}(\mathbf{a}),\,\boldsymbol{\mu}(\mathbf{a})},$$

where $\mathcal{L}$ is the Lagrange function,

$$\mathcal{L}(\mathbf{x},\boldsymbol{\mu};\mathbf{a})=f(\mathbf{x};\mathbf{a})-\boldsymbol{\mu}^{\mathrm{T}}\mathbf{h}(\mathbf{x};\mathbf{a}).$$

Proof The proof will parallel the equality constraint case. By chain rule, as in the previous Theorem,

$$\nabla_{\mathbf{a}}M(\mathbf{a})^{\mathrm{T}}=\nabla_{\mathbf{x}}f(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a})+\nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})\Big|_{\mathbf{x}(\mathbf{a})}. \quad (4)$$

Since $\mathbf{x}(\mathbf{a})$ is a local optimal solution, from the first-order necessary condition that

$$\begin{aligned}\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}(\mathbf{a}),\boldsymbol{\lambda}(\mathbf{a});\mathbf{a})&=\nabla_{\mathbf{x}}f(\mathbf{x}(\mathbf{a});\mathbf{a})-\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\boldsymbol{\mu}(\mathbf{a})\\&=\mathbf{0},\end{aligned}$$

we have,

$$\nabla_{\mathbf{x}}f(\mathbf{x}(\mathbf{a});\mathbf{a})=\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\boldsymbol{\mu}(\mathbf{a}). \quad (5)$$

Now assume that the first b inequality constraints are binding and the rest are not. Since $\mathbf{x}(\mathbf{a})$ is a $\boldsymbol{C}^1$ function, when $\mathbf{a}$ changes sufficiently small, there will be a corresponding change in $\mathbf{x}(\mathbf{a})$ such that the first r inequality constraints are still binding and the rest are not. Thus, the gradient of the first r inequality constraints with respect to $\mathbf{a}$ will be zero as in the equality constraint case, and the multipliers of the last $m-r$ constraints are zeros. That is, if

$$\mathbf{h}^r(\mathbf{x}(\mathbf{a});\mathbf{a}) = \begin{bmatrix} h^1((\mathbf{x}(\mathbf{a});\mathbf{a}) \\ \vdots \\ h^r((\mathbf{x}(\mathbf{a});\mathbf{a}) \end{bmatrix} = \mathbf{0},$$

with associated Lagrange multiplier $\boldsymbol{\mu}^r(\mathbf{a})$ then

$$\begin{aligned} \nabla_{\mathbf{a}}\mathbf{h}^r(\mathbf{x}(\mathbf{a});\mathbf{a}) &= \nabla_{\mathbf{x}}\mathbf{h}^r(\mathbf{x}(\mathbf{a});\mathbf{a})\nabla\mathbf{x}(\mathbf{a}) + \nabla_{\mathbf{a}}\mathbf{h}^r(\mathbf{x};\mathbf{a})\Big|_{\mathbf{x}(\mathbf{a})} \\ &= \mathbf{0} \end{aligned}$$

and so,

$$\nabla_{\mathbf{x}}\mathbf{h}^r(\mathbf{x}(\mathbf{a});\mathbf{a})\nabla\mathbf{x}(\mathbf{a}) = -\nabla_{\mathbf{a}}\mathbf{h}^r(\mathbf{x};\mathbf{a}) \qquad (6)$$

From (4), (5) and (6)

$$\begin{aligned} \nabla_{\mathbf{a}}M(\mathbf{a})^{\mathrm{T}} &= \boldsymbol{\mu}(\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) + \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a})} \\ &= \boldsymbol{\mu}^r(\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{x}}\mathbf{h}^r(\mathbf{x}(\mathbf{a});\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{a}}\mathbf{x}(\mathbf{a}) + \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a})} \\ &= \boldsymbol{\mu}^r(\mathbf{a})^{\mathrm{T}}\left(-\nabla_{\mathbf{a}}\mathbf{h}^r(\mathbf{x};\mathbf{a})\Big|_{\mathbf{x}(\mathbf{a})}\right) + \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a})} \\ &= \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a})} - \boldsymbol{\mu}^r(\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{a}}\mathbf{h}^r(\mathbf{x};\mathbf{a})\Big|_{\mathbf{x}(\mathbf{a})} \\ &= \nabla_{\mathbf{a}}f(\mathbf{x};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a})} - \boldsymbol{\mu}(\mathbf{a})^{\mathrm{T}}\nabla_{\mathbf{a}}\mathbf{h}(\mathbf{x};\mathbf{a})\Big|_{\mathbf{x}(\mathbf{a})} \\ &= \nabla_{\mathbf{a}}\mathcal{L}(\mathbf{x},\boldsymbol{\mu};\mathbf{a})^{\mathrm{T}}\Big|_{\mathbf{x}(\mathbf{a}),\,\boldsymbol{\mu}(\mathbf{a})}. \end{aligned}$$

The second and the fifth equalities utilize the fact that $\boldsymbol{\mu}(\mathbf{a}) = \begin{bmatrix}\boldsymbol{\mu}^r(\mathbf{a}) \\ \mathbf{0}\end{bmatrix}$.❑

HW Show that the theorem on the meaning of the multipliers in the inequality constraint case is just a special case of this Envelope Theorem in the inequality constraint case.

HW Write and prove the Envelope Theorem for optimization under mixed constraints.