On the (Non-)Differentiability of the Optimal Value Function When the Optimal Solution Is Unique*

Daisuke Oyama

Faculty of Economics, University of Tokyo
Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
oyama@e.u-tokyo.ac.jp

and

Tomoyuki Takenawa

Faculty of Marine Technology,
Tokyo University of Marine Science and Technology
2-1-6 Etchu-jima, Koto-ku, Tokyo 135-8533, Japan
takenawa@kaiyodai.ac.jp

December 25, 2013; revised January 30, 2018

Abstract

We present examples of a parameterized optimization problem, with a continuous objective function differentiable with respect to the parameter, that admits a unique optimal solution, but whose optimal value function is not differentiable. We also show independence of Danskin's and Milgrom and Segal's envelope theorems. *Journal of Economic Literature* Classification Number: C61.

Keywords: parameterized optimization; sensitivity analysis; value function; envelope theorem.

*We would like to thank Atsushi Kajii, Kazuya Kamiya, Jan Werner, and anonymous referees for helpful comments and discussions.

1 Introduction

Parameterized optimization problems are ubiquitous in economics, from classical price theory to dynamic macroeconomics, game theory, mechanism design, and so on. There, the envelope theorem serves as a standard tool in understanding the marginal effects of changes in the parameter, such as price or technology, on the value of the optimal choice of the agents in the model. While textbook envelope theorems usually only derive a formula ("envelope formula") that the derivative of the value function, the optimal value as a function of the parameter, should satisfy under the (often implicit) assumption that the value function is differentiable, a rigorous statement of the theorem also describes a sufficient condition on the primitives under which this assumption holds true. The latter issue, the differentiability of the value function, is what we are concerned with in this paper.

We consider the following setting. Let X be a nonempty topological space (the choice set), and $A \subset \mathbb{R}$ a nonempty open set (the parameter space). The objective function $f \colon X \times A \to \mathbb{R}$ is to be maximized with respect to $x \in X$, given $\alpha \in A$. The optimal value function is given by

$$v(\alpha) = \sup_{x \in X} f(x, \alpha),$$

associated with the optimal solution correspondence

$$X^*(\alpha) = \{x \in X \mid f(x, \alpha) = v(\alpha)\},$$

where we assume that the partial derivative of f with respect to α, f_α, exists and that $X^*(\alpha) \neq \emptyset$ for all $\alpha \in A$. We are interested in the differentiability of the value function v (in the classical sense, rather than notions such as directional differentiability or subdifferentiability).

1. *If v is assumed to be differentiable* at $\alpha = \bar{\alpha}$, then it is easy to derive the envelope formula: for any $\bar{x} \in X^*(\bar{\alpha})$,

   $$v'(\bar{\alpha}) = f_\alpha(\bar{x}, \bar{\alpha}).$$

   Indeed, fix any $\bar{x} \in X^*(\bar{\alpha})$. Then the function $g(\alpha) = f(\bar{x}, \alpha) - v(\alpha)$, which is differentiable at $\bar{\alpha}$, is maximized at $\bar{\alpha}$, so the first-order condition $g'(\bar{\alpha}) = 0$ yields the formula.

2. One can easily construct an example in which the value function is not differentiable *when there are more than one solutions*. For example, let $X = \mathbb{R}$ and $A = (-1, 1)$, and consider

   $$f(x, \alpha) = -\frac{1}{4}x^4 - \frac{\alpha}{3}x^3 + \frac{1}{2}x^2 + \alpha x - \frac{1}{4},$$

with $f_x(x, \alpha) = -(x+1)(x+\alpha)(x-1)$. Then we have

$$v(\alpha) = \frac{2}{3}|\alpha|, \quad X^*(\alpha) = \begin{cases} \{-1\} & \text{if } \alpha < 0, \\ \{-1, 1\} & \text{if } \alpha = 0, \\ \{1\} & \text{if } \alpha > 0, \end{cases}$$

where v is not differentiable at $\alpha = 0$, for which there are two optimal solutions.

3. The question we ask in this paper is: is the value function always differentiable *when the optimal solution is unique* (and the objective function is continuous and the solution correspondence admits a continuous selection)? The answer is no: we present in Section 2 an example in which f_α exists and X^* is a single-valued continuous function, but nevertheless v is not differentiable at some $\bar{\alpha}$ (Example 2.1). The main feature in our example is that f_α is not continuous at $(x, \alpha) = (\bar{x}, \bar{\alpha})$, where $\{\bar{x}\} = X^*(\bar{\alpha})$. In fact, if X^* admits a selection continuous at $\bar{\alpha}$ and f_α is continuous jointly in (x, α) at $(\bar{x}, \bar{\alpha})$, then v must be differentiable at $\bar{\alpha}$ (Proposition 2.1).

Numerous forms of sufficient conditions for the differentiability of the value function have been obtained in the literature. In Section 3, we discuss the results by Danskin (1966, 1967) and Milgrom and Segal (2002). Danskin's theorem also assumes the continuity of f_α, and when applied to the case where the optimal solution is unique, his assumptions are slightly stronger than those in Proposition 2.1 mentioned above, while they are not nested in general. Our Example 2.1 illustrates that the continuity of f_α is indispensable also in Danskin's theorem.

Milgrom and Segal (2002, Theorem 3) provide a sufficient condition in terms of the equidifferentiability of the objective function f. It turns out that our example does not satisfy this condition. We present examples that illustrate that neither of the continuity of f_α at $(\bar{x}, \bar{\alpha})$ and the equidifferentiability of $\{f(x, \cdot)\}_{x \in X}$ implies the other, showing that the conditions in Danskin's theorem, or our Proposition 2.1, and those in Milgrom and Segal's theorem are independent from each other.

It has been known that certain convexity/concavity conditions allow the differentiability of the value function. For instance, the support function of a closed convex set in a finite-dimensional space, examples of which include the profit, cost, and expenditure functions in price theory, is differentiable if and only if the maximum (or minimum) is attained at a single point; see Mas-Colell et al. (1995, Proposition 3.F.1) or Rockafellar (1970, Corollary 25.1.3). In this case, the partial derivative of the objective function with respect to the parameter is clearly continuous. In fact, we show in Section 5.1 that this theorem, well known from convex analysis, also follows from (a multidimensional parameter version of) our envelope theorem Proposition 2.1, despite the possible unboundedness of the choice set. If the objective function f is concave jointly in the choice variable x and the parameter α and if f_α exists, then the value function, which is necessarily concave, is always differentiable; see Hogan (1973), Benveniste and Scheinkman (1979), or Milgrom and Segal (2002, Corollary 3). For this result, topology is not needed for the choice set (only being a convex subset of a linear space), and hence the continuity of f_α is irrelevant.

In Section 4, we extend our example to optimization problems with inequality constraints that vary with the parameter α. We provide examples with a binding constraint in which the Lagrange function L is differentiable in α and the optimal solution and the Kuhn-Tucker vector (which constitute a saddle point of L) are unique and continuous in α, but the optimal value function is not differentiable. Again, in these examples, L_α fails to be continuous: in fact, if a function $L(x, y, \alpha)$ has a saddle point selection $(\bar{x}(\alpha), \bar{y}(\alpha))$ that is continuous in α and L_α is continuous in (x, y, α), then its saddle value function $L(\bar{x}(\alpha), \bar{y}(\alpha), \alpha)$ is differentiable in α (Proposition 4.1). We also observe that if the value function of the constrained problem is concave, then the existence of a continuous selection of Kuhn-Tucker vectors is sufficient for the differentiability (Proposition 4.3).[1]

In Section 5, we apply our analysis within the context of price theory. Section 5.1 considers the profit function (or the support function) for a closed convex production set, where, as mentioned, we present a proof for its differentiability that uses our Proposition 2.1. Section 5.2 concerns the value functions for consumption choice. We first state envelope theorems derived from Proposition 2.1 for the indirect utility function and the expenditure function: if the utility function is continuous and locally nonsatiated and has partial derivatives which are continuous and nonvanishing at the optimum, then these functions are differentiable whenever the optimal solutions are unique. Then, extending our main example into this framework, we construct an example of a continuous utility function with positive partial derivatives for which, for some fixed price vector, the Walrasian and Hicksian demands are unique and continuous in wealth w and required utility u, but the indirect utility and expenditure functions fail to be differentiable in w and u, respectively. In this example, the partial derivatives of the utility function are not continuous, demonstrating that the continuous differentiability condition cannot be replaced with the mere existence of partial derivatives for the differentiability of the value functions in this case as well.

[1] This result extends Corollary 3 in Milgrom and Segal (2002) to the case of parametric constraints. See also Marimon and Werner (2016).

2 (Non-)Differentiability of the Value Function

Let X be a nonempty topological space, and $A \subset \mathbb{R}$ a nonempty open set. Given the objective function $f\colon X \times A \to \mathbb{R}$, we consider the optimal value function

$$v(\alpha) = \sup_{x \in X} f(x, \alpha),$$

associated with the optimal solution correspondence

$$X^*(\alpha) = \{x \in X \mid f(x, \alpha) = v(\alpha)\}.$$

We are interested in the differentiability of v when X^* is point-valued. We first state a sufficient condition of direct relevance for our study.

Proposition 2.1. *Assume that*

(a) *X^* has a selection x^* continuous at $\bar{\alpha}$, and*

(b) *f is differentiable in α in a neighborhood of $(x^*(\bar{\alpha}), \bar{\alpha})$, and f_α is continuous in (x, α) at $(x^*(\bar{\alpha}), \bar{\alpha})$.*

Then v is differentiable at $\bar{\alpha}$ with $v'(\bar{\alpha}) = f_\alpha(x^*(\bar{\alpha}), \bar{\alpha})$.

Assumption (a) holds if X^* is nonempty-valued and upper semi-continuous (which holds true, e.g., when f is continuous and X is compact) and $X^*(\bar{\alpha})$ is a singleton, in which case any selection of X^* is continuous at $\bar{\alpha}$.

A version of this proposition is found in the lecture notes by Border (2015, Corollary 299). For completeness, we present the proof in Appendix A.1.

Our main observation in this paper is that, even when there is a unique optimal solution, the differentiability of v may fail if one drops the continuity of f_α.

Proposition 2.2. *There exists a continuous function $f\colon X \times A \to \mathbb{R}$ such that*

(a) *$X^*(\alpha)$ is a singleton for all α and is continuous in α (as a single-valued function), and*

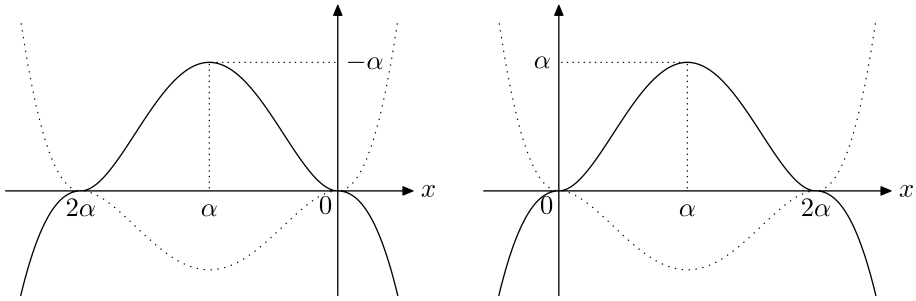
(b) *f is differentiable in α,*

but v is not differentiable at some α.

In the following, we present an example of such a function f.

Example 2.1. Let $X = \mathbb{R}$ and $A = \mathbb{R}$, and let the continuous function f be defined by

$$f(x, \alpha) = \begin{cases} -\dfrac{1}{\alpha^3} x^2 (x - 2\alpha)^2 & \text{if } 2\alpha < x < 0, \\ \dfrac{1}{\alpha^3} x^2 (x - 2\alpha)^2 & \text{if } 0 < x < 2\alpha, \\ -x^2 (x - 2\alpha)^2 & \text{otherwise.} \end{cases}$$



(1) $\alpha < 0$ (2) $\alpha > 0$

Figure 1: Graph of $f(\cdot, \alpha)$

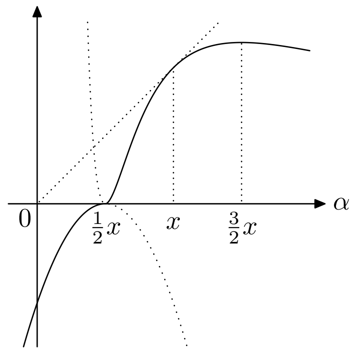


Figure 2: Graph of $f(x, \cdot)$ $(x > 0)$

The left and the right panels in Figure 1 depict the graph of $f(\cdot, \alpha)$ for $\alpha < 0$ and $\alpha > 0$, respectively, while Figure 2 depicts the graph of $f(x, \cdot)$ for $x > 0$.

The function f is differentiable in α at all (x, α) with

$$f_\alpha(x, \alpha) = \begin{cases} \dfrac{1}{\alpha^4} x^2(x - 2\alpha)(3x - 2\alpha) & \text{if } 2\alpha < x < 0, \\ -\dfrac{1}{\alpha^4} x^2(x - 2\alpha)(3x - 2\alpha) & \text{if } 0 < x < 2\alpha, \\ 4x^2(x - 2\alpha) & \text{otherwise.} \end{cases}$$

One can verify that

$$X^*(\alpha) = \{\alpha\},$$

and

$$v(\alpha) = |\alpha|,$$

which is not differentiable at $\alpha = 0$.

While continuous separately in x and in α, f_α is not continuous jointly in (x, α) at $(x, \alpha) = (0, 0)$. To see this, set $x = c\alpha$, $\alpha \neq 0$, and let $\alpha \to 0$.

Then, depending on c we have

$$\lim_{\alpha \to 0-} f_\alpha(c\alpha, \alpha) = c^2(c-2)(3c-2),$$
$$\lim_{\alpha \to 0+} f_\alpha(c\alpha, \alpha) = -c^2(c-2)(3c-2),$$

so that $0 \neq \lim_{\alpha \to 0-} f_\alpha(c\alpha, \alpha) \neq \lim_{\alpha \to 0+} f_\alpha(c\alpha, \alpha) \neq 0$ if $0 < c < 2$, $c \neq 2/3$, while

$$\lim_{\alpha \to 0-} f_\alpha(c\alpha, \alpha) = \lim_{\alpha \to 0+} f_\alpha(c\alpha, \alpha) = 0$$

otherwise (and $\lim_{x \to 0} f_\alpha(x, 0) = 0$). Thus, Assumption (b) in Proposition 2.1 is violated. □

Remark 2.1. It is easy to find an example of a non-differentiable value function where $X^*(\alpha)$ is a singleton for all α and f is differentiable in α, but f or X^* (as a single-valued function) is not continuous. As such an example, let $X = \mathbb{R}$ and $A = \mathbb{R}$, and let f be defined by

$$f(x, \alpha) = \begin{cases} \alpha - (x-\alpha)^2 & \text{if } x \geq 0, \\ -\alpha - (x-\alpha)^2 & \text{if } x < 0. \end{cases}$$

Then $X^*(\alpha) = \{\alpha\}$, and $v(\alpha) = |\alpha|$, which is not differentiable at $\alpha = 0$. Here, f is not continuous and X^* is continuous with the Euclidean topology on X, while if X is endowed with the discrete topology, then f is continuous but X^* is not.

3 Relation to Danskin and Milgrom and Segal

In this section, we discuss the relation to the sufficient conditions for (directional) differentiability obtained by Danskin (1966) and Milgrom and Segal (2002) and between them.

3.1 Danskin (1966)

A careful inspection of the proof of Theorem 1 in Danskin (1966), or the earlier result in Gross (1954),[2] shows the following to hold.

Theorem 3.1 (Danskin)**.** *Assume that*

(a) *X^* is upper semi-continuous with $X^*(\alpha) \neq \emptyset$ for all $\alpha \in A$, and $X^*(\bar{\alpha})$ is compact, and*

(b) *for all $\bar{x} \in X^*(\bar{\alpha})$, f is differentiable in α in a neighborhood of $(\bar{x}, \bar{\alpha})$, and f_α is continuous in (x, α) at $(\bar{x}, \bar{\alpha})$.*

[2] See also the Addendum in Danskin (1966).

Then v is left- and right-hand differentiable at $\bar{\alpha}$, and the directional derivatives of v at $\bar{\alpha}$ are given by

$$v'(\bar{\alpha}+) = \max_{\bar{x} \in X^*(\bar{\alpha})} f_\alpha(\bar{x}, \bar{\alpha}), \tag{1}$$

$$v'(\bar{\alpha}-) = \min_{\bar{x} \in X^*(\bar{\alpha})} f_\alpha(\bar{x}, \bar{\alpha}). \tag{2}$$

v is differentiable at $\bar{\alpha}$ if and only if $f_\alpha(\bar{x}, \bar{\alpha}) = f_\alpha(\bar{x}', \bar{\alpha})$ for all $\bar{x}, \bar{x}' \in X^*(\bar{\alpha})$, in which case $v'(\bar{\alpha}) = f_\alpha(\bar{x}, \bar{\alpha})$ for any $\bar{x} \in X^*(\bar{\alpha})$.

Two stronger assumptions are made in the original statement in Danskin (1966, Theorem 1) which are relaxed here: first, it is assumed that X is compact and f is continuous, while its consequence, Assumption (a), is what is used; second, the (existence and) continuity of f_α is assumed on the whole space of $X \times A$, while the continuity only on $X^*(\bar{\alpha}) \times \{\bar{\alpha}\}$ is sufficient for the conclusion. For completeness, we present the proof in Appendix A.2.

When $X^*(\bar{\alpha})$ is a singleton, under Assumption (a) any selection of X^* is continuous at $\bar{\alpha}$, so that the differentiability of v at $\bar{\alpha}$ follows from Proposition 2.1, while in general, the assumptions in Proposition 2.1 and Theorem 3.1 are not nested.

Our Example 2.1 in the previous section illustrates that the continuity of f_α is indispensable for Theorem 3.1. In Example 2.1, f_α is not continuous at $(x, \alpha) = (0, 0)$, where $X^*(0) = \{0\}$. The value function v is left- and right-hand differentiable at $\alpha = 0$, while the directional derivatives do not obey the formulas (1) and (2): $v'(0+) = 1$ and $v'(0-) = -1$, while $f_\alpha(0, 0) = 0$.

3.2 Milgrom and Segal (2002)

Milgrom and Segal (2002, Theorem 3) provide a sufficient condition in terms of equidifferentiability. A family of functions $\{f(x, \cdot)\}_{x \in X}$ is equidifferentiable at $\bar{\alpha} \in A$ if for all $x \in X$, $f(x, \cdot)$ is differentiable in α at $\bar{\alpha}$, and for any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in X$,

$$\left| \frac{f(x, \alpha) - f(x, \bar{\alpha})}{\alpha - \bar{\alpha}} - f_\alpha(x, \bar{\alpha}) \right| \leq \varepsilon$$

whenever $|\alpha - \bar{\alpha}| \leq \delta$, $\alpha \neq \bar{\alpha}$. Note that in their theorem, no structure is imposed on the choice set X.

Theorem 3.2 (Milgrom and Segal)**.** *Assume that*

(a) *$X^*(\alpha) \neq \emptyset$ for all $\alpha \in A$, and*

(b) *$\{f(x, \cdot)\}_{x \in X}$ is equidifferentiable at $\bar{\alpha}$, and* $\sup_{x \in X} |f_\alpha(x, \bar{\alpha})| < \infty$.

Then v is left- and right-hand differentiable at $\bar{\alpha}$, and for any selection x^* of X^*, the directional derivatives of v at $\bar{\alpha}$ are given by

$$v'(\bar{\alpha}+) = \lim_{\alpha \to \bar{\alpha}+} f_\alpha(x^*(\alpha), \bar{\alpha}), \tag{3}$$

$$v'(\bar{\alpha}-) = \lim_{\alpha \to \bar{\alpha}-} f_\alpha(x^*(\alpha), \bar{\alpha}). \tag{4}$$

v is differentiable at $\bar{\alpha}$ if and only if $f_\alpha(x^*(\alpha), \bar{\alpha})$ is continuous in α at $\bar{\alpha}$ for some selection x^* of X^*, in which case $v'(\bar{\alpha}) = f_\alpha(\bar{x}, \bar{\alpha})$ for any $\bar{x} \in X^*(\bar{\alpha})$.

If one assumed in Assumption (b) in Proposition 2.1 or Theorem 3.1 that X is compact and f_α exists and is continuous on $X \times A$,[3] then Assumption (b) in Theorem 3.2 would hold (see Corollary 4 in Milgrom and Segal (2002)), and hence the conclusion of Proposition 2.1 or Theorem 3.1 would follow from Theorem 3.2. Otherwise, the assumptions in Proposition 2.1 or Theorem 3.1 and those in Theorem 3.2 are not nested; see Subsection 3.3.

In our Example 2.1, if we restrict X to be any bounded interval containing $x = 0$, all the conditions in Theorem 3.2 except the equidifferentiability are satisfied. To see that the equidifferentiability fails at $\alpha = 0$, for any $\delta > 0$ let $(x, \alpha) = (\delta, \delta)$. Then we have

$$\left| \frac{f(\delta, \delta) - f(\delta, 0)}{\delta - 0} - f_\alpha(\delta, 0) \right| = \left| \frac{\delta - (-4\delta^4)}{\delta} - 4\delta^3 \right| = 1$$

for all $\delta > 0$. Again, the directional derivatives of v at $\alpha = 0$ do not obey the formulas (3) and (4): $v'(0+) = 1$ and $v'(0-) = -1$, while $f_\alpha(x^*(\alpha), 0) = 4|\alpha|^3 \to 0$ as $\alpha \to 0$.

We note that if $f(x, \alpha)$ is continuous in x for all $\alpha \in A$, then the equidifferentiability of $\{f(x, \cdot)\}_{x \in X}$ at $\bar{\alpha}$ implies that $f_\alpha(x, \bar{\alpha})$ is continuous in x (but it may not be continuous in α; see Example 3.1 in Subsection 3.3).

Proposition 3.3. *Assume that $f(\cdot, \alpha)$ is continuous for all $\alpha \in A$. If $\{f(x, \cdot)\}_{x \in X}$ is equidifferentiable at $\bar{\alpha}$, then $f_\alpha(\cdot, \bar{\alpha})$ is continuous in x.*

The proof is given in Appendix A.3.

Thus, if $f(\cdot, \alpha)$ is continuous for all $\alpha \in A$, X^* has a selection that is continuous at $\bar{\alpha}$, and Assumption (b) in Theorem 3.2 holds, then v is differentiable at $\bar{\alpha}$ by Theorem 3.2.

3.3 Independence of the Theorems

In this section, we see the independence of the assumptions in Proposition 2.1 or Theorem 3.1 and those in Theorem 3.2. First, the equidifferentiability of f does not imply the continuity of f_α, and thus Theorem 3.2 does

[3] Or more weakly, that for some neighborhood B of $\bar{\alpha}$ in A, $X^*(B)$ is relatively compact (i.e., its closure $\mathrm{cl}(X^*(B))$ is compact) and f_α exists and is continuous on $\mathrm{cl}(X^*(B)) \times B$.

not follow from Proposition 2.1 or Theorem 3.1. One can easily construct a continuous function $f\colon X \times A \to \mathbb{R}$ such that for some $\bar{\alpha} \in A$, the assumptions in Theorem 3.2 hold, but f_α is not continuous at $(\bar{x}, \bar{\alpha})$ for any $\bar{x} \in X^*(\bar{\alpha})$.

Example 3.1. Let $X = \mathbb{R}$ and $A = \mathbb{R}$, and let $f\colon X \times A \to \mathbb{R}$ be defined by

$$f(x, \alpha) = \begin{cases} -x^2 + \alpha^2 \sin \dfrac{1}{\alpha} & \text{if } \alpha \neq 0, \\ -x^2 & \text{if } \alpha = 0. \end{cases}$$

This function is differentiable at all (x, α), whose derivative with respect to α is given by

$$f_\alpha(x, \alpha) = \begin{cases} 2\alpha^2 \sin \dfrac{1}{\alpha} - \cos \dfrac{1}{\alpha} & \text{if } \alpha \neq 0, \\ 0 & \text{if } \alpha = 0, \end{cases}$$

which is discontinuous in α at $\alpha = 0$. It is immediate that $X^*(\alpha) = \{0\}$ for all α, and the value function $v(\alpha) = f(0, \alpha)$ is differentiable with $v'(\alpha) = f_\alpha(0, \alpha)$ for all α. Clearly, the convergence $(f(x, \alpha') - f(x, \alpha))/(\alpha' - \alpha) \to f_\alpha(x, \alpha)$ as $\alpha' \to \alpha$ does not depend on x, so that $\{f(x, \cdot)\}_{x \in X}$ is equidifferentiable at all α including $\alpha = 0$. ▯

Second, the continuity of f_α does not imply the equidifferentiability of f, and thus neither Proposition 2.1 nor Theorem 3.1 follows from Theorem 3.2.

Proposition 3.4. *There exists a continuous function $f\colon X \times A \to \mathbb{R}$ such that for some $\bar{\alpha} \in A$, the assumptions in Proposition 2.1 and Theorem 3.1 hold, but $\{f(x, \cdot)\}_{x \in U}$ is not equidifferentiable at $\bar{\alpha}$ for any neighborhood U of $X^*(\bar{\alpha})$.*

This is proved by the following example.

Example 3.2. Let $X = \mathbb{R}$ and $A = \mathbb{R}$, and let $f\colon X \times A \to \mathbb{R}$ be defined by

$$f(x, \alpha) = \begin{cases} -x^2 + g_k(x, \alpha) & \text{if } \dfrac{1}{2^k} < x < \dfrac{1}{2^k}\left[1 + \dfrac{2}{\pi} \tan^{-1} \alpha\right], \\ -x^2 & \text{otherwise}, \end{cases}$$

where k runs through all the integers, and g_k is defined by

$$g_k(x, \alpha) = \frac{2^{(2+\gamma)k}}{(\tan^{-1} \alpha)^3} \left(x - \frac{1}{2^k}\right)^2 \left(x - \frac{1}{2^k}\left[1 + \frac{2}{\pi} \tan^{-1} \alpha\right]\right)^2$$

with some constant $\gamma \in (0,2)$. Since $\lim_{x \to \pm\infty} f(x,\alpha) = -\infty$ and hence we can restrict x to a compact set, and since f is continuous, X^* is nonempty-valued and upper semi-continuous. Since $X^*(0) = \{0\}$, Assumption (a) in Proposition 2.1 as well as Assumption (a) in Theorem 3.1 hold with $\bar{\alpha} = 0$.[4]

The partial derivative of g_k with respect to α is calculated as

$$\begin{aligned}\frac{\partial g_k}{\partial \alpha}(x,\alpha) = &-\frac{2^{(2+\gamma)k}}{(1+\alpha^2)(\tan^{-1}\alpha)^4}\left(x - \frac{1}{2^k}\right)^2 \\ &\times \left(x - \frac{1}{2^k}\left[1 + \frac{2}{\pi}\tan^{-1}\alpha\right]\right)\left(3x - \frac{1}{2^k}\left[3 + \frac{2}{\pi}\tan^{-1}\alpha\right]\right)\end{aligned}$$

(where the formula $(\tan^{-1}\alpha)' = 1/(1+\alpha^2)$ is used). To verify that Assumption (b) in Proposition 2.1 or Theorem 3.1 is satisfied, note that any point (x,α) such that $\frac{1}{2^k} < x < \frac{1}{2^k}\left[1 + \frac{2}{\pi}\tan^{-1}\alpha\right]$ is written as

$$(x,\alpha) = \left(\frac{1}{2^k}\left[1 + c\frac{2}{\pi}\tan^{-1}\alpha\right], \alpha\right)$$

for some $c \in (0,1)$. For such points (x,α), we have

$$f_\alpha(x,\alpha) = \frac{\partial g_k}{\partial \alpha}(x,\alpha) = -\left(\frac{2}{\pi}\right)^4 \frac{1}{2^{(2-\gamma)k}(1+\alpha^2)} c^2(c-1)(3c-1).$$

First, since this tends to 0 as $c \to 1$, it follows that f is differentiable in α at points (x,α) such that $x = \frac{1}{2^k}\left[1 + \frac{2}{\pi}\tan^{-1}\alpha\right]$ (the differentiability at other points is clear). Second, since $k \to \infty$ as $x \to 0$, for those points (x,α) we have $f_\alpha(x,\alpha) \to 0$ as $(x,\alpha) \to (0,0)$, which implies that f_α is continuous in (x,α) at $(x,\alpha) = (0,0)$ (while it is discontinuous at points $(x,\alpha) = (1/2^k, 0)$).

Now we claim that for any neighborhood B of $x = 0$, $\{f(x,\cdot)\}_{x \in B}$ is not equidifferentiable at $\alpha = 0$. Let $B = [-\eta, \eta]$, $\eta > 0$, and set, for example, $\varepsilon = 1/(\pi^4 2^{(2-\gamma)K+1})$, where K is such that $1/2^{K-1} \leq \eta$. For any $\delta > 0$, let $\alpha' \in (0,\delta]$ be such that $(\tan^{-1}\alpha')/\alpha' \geq 1/2$ (note that $(\tan^{-1}\alpha)/\alpha \to 1$ as $\alpha \to 0$), and let $x' = \frac{1}{2^K}\left[1 + \frac{1}{\pi}\tan^{-1}\alpha'\right] \in B$. Then we have

$$\frac{f(x',\alpha') - f(x',0)}{\alpha' - 0} - f_\alpha(x',0) = \frac{1}{\pi^4 2^{(2-\gamma)K}} \frac{\tan^{-1}\alpha'}{\alpha'} \geq \varepsilon,$$

as claimed. □

4 Parameterized Constraints

In this section, we discuss optimization problems with parameterized constraints and provide examples analogous to Example 2.1.

[4] The condition $\gamma > 0$ implies that $0 \notin X^*(\alpha)$ for $\alpha > 0$, while $\gamma < 2$ will ensure the continuity of f_α at $(x,\alpha) = (0,0)$. Let $\gamma < 1$ if one wants f to be differentiable everywhere with respect to x as well.

4.1 Saddle Point Problems

We first consider parameterized saddle point problems. Let X and Y be nonempty topological spaces, and $A \subset \mathbb{R}$ a nonempty open set. For $L\colon X \times Y \times A \to \mathbb{R}$, $(\bar{x}, \bar{y}) \in X \times Y$ is a *saddle point* of L at $\alpha \in A$ if

$$L(x, \bar{y}, \alpha) \leq L(\bar{x}, \bar{y}, \alpha) \leq L(\bar{x}, y, \alpha)$$

for all $x \in X$ and all $y \in Y$. It is well known that if the set of saddle points of L at α is nonempty, then it is written as the product set $\bar{X}(\alpha) \times \bar{Y}(\alpha)$, where

$$\bar{X}(\alpha) = \left\{ x \in X \;\middle|\; \inf_{y \in Y} L(x, y, \alpha) = \sup_{x' \in X} \inf_{y \in Y} L(x', y, \alpha) \right\},$$
$$\bar{Y}(\alpha) = \left\{ y \in Y \;\middle|\; \sup_{x \in X} L(x, y, \alpha) = \inf_{y' \in Y} \sup_{x \in X} L(x, y', \alpha) \right\},$$

and for any $(\bar{x}, \bar{y}) \in \bar{X}(\alpha) \times \bar{Y}(\alpha)$, we have

$$L(\bar{x}, \bar{y}, \alpha) = \inf_{y \in Y} L(\bar{x}, y, \alpha) = \sup_{x \in X} L(x, \bar{y}, \alpha).$$

We denote this value, the *saddle value* of L at α, by $\bar{v}(\alpha)$.

As a counterpart of Proposition 2.1, we have the following (see also Border (2015, Theorem 298)).

Proposition 4.1. *Assume that*

(a) $\bar{X} \times \bar{Y}$ *has a selection* $(\bar{x}, \bar{y})$ *continuous at* $\bar{\alpha}$, *and*

(b) L *is differentiable in* α *in a neighborhood of* $(\bar{x}(\bar{\alpha}), \bar{y}(\bar{\alpha}), \bar{\alpha})$, *and* L_α *is continuous in* (x, y, α) *at* $(\bar{x}(\bar{\alpha}), \bar{y}(\bar{\alpha}), \bar{\alpha})$.

Then $\bar{v}$ *is differentiable at* $\bar{\alpha}$ *with* $\bar{v}'(\bar{\alpha}) = L_\alpha(\bar{x}(\bar{\alpha}), \bar{y}(\bar{\alpha}), \bar{\alpha})$.

The proof of this proposition is similar to that of Proposition 2.1 and thus is omitted.

4.2 Constrained Optimization Problems

For functions $f\colon X \times A \to \mathbb{R}$ and $g\colon X \times A \to \mathbb{R}^k$, consider the maximization problem with parameterized inequality constrains:

$$v(\alpha) = \sup_{x \in X(\alpha)} f(x, \alpha),$$
$$X^*(\alpha) = \{x \in X \mid x \in X(\alpha),\ f(x, \alpha) = v(\alpha)\},$$

where

$$X(\alpha) = \{x \in X \mid g(x, \alpha) \geq 0\}.$$

Let $L\colon X \times \mathbb{R}^k_+ \times A \to \mathbb{R}$ be the associated Lagrange function:

$$L(x, y, \alpha) = f(x, \alpha) + \sum_{i=1}^{k} y_i g_i(x, \alpha).$$

Let

$$Y^*(\alpha) = \left\{ y \in \mathbb{R}^k_+ \;\middle|\; \sup_{x \in X} L(x, y, \alpha) = v(\alpha) \right\},$$

which is the set of *Kuhn-Tucker vectors* (Rockafellar (1970)) at α. Note that $(x^*, y^*) \in X^*(\alpha) \times Y^*(\alpha)$ if and only if it is a saddle point of L at α, and in this case, $v(\alpha) = L(x^*, y^*, \alpha)$.

We show that even if the optimal solution and the Kuhn-Tucker vector are unique, the optimal value function may not be differentiable.

Proposition 4.2. *There exist continuous functions $f\colon X \times A \to \mathbb{R}$ and $g\colon X \times A \to \mathbb{R}^k$ such that*

(a) *$X^*(\alpha) \times Y^*(\alpha)$ is a singleton for all α and is continuous in α (as a single-valued function), where the element of $Y^*(\alpha)$ is nonzero for all α, and*

(b) *f and g are differentiable in α,*

but v is not differentiable at some α.

Modifying Example 2.1, we construct two examples of such functions f and g. In the former (Example 4.1), f_α fails to be continuous, violating Assumption (b) in Proposition 4.1, while g_α does in the latter (Example 4.2).[5]

Example 4.1. Let $X = \mathbb{R}$, $A = \mathbb{R}$, and $k = 1$, and let continuous functions f and g be defined by

$$\begin{aligned} f(x, \alpha) &= f^0(x, \alpha) + x, \\ g(x, \alpha) &= -x + \alpha, \end{aligned}$$

where f^0 is the function defined in Example 2.1, i.e.,

$$f^0(x, \alpha) = \begin{cases} -\dfrac{1}{\alpha^3} x^2 (x - 2\alpha)^2 & \text{if } 2\alpha < x < 0, \\ \dfrac{1}{\alpha^3} x^2 (x - 2\alpha)^2 & \text{if } 0 < x < 2\alpha, \\ -x^2 (x - 2\alpha)^2 & \text{otherwise,} \end{cases}$$

which is uniquely maximized at $x = \alpha$ (on X). One can verify that

$$X^*(\alpha) = \{\alpha\},$$

[5]These examples also demonstrate that the continuity of L_α is indispensable in Theorem 5 in Milgrom and Segal (2002) and Theorem 3 in Marimon and Werner (2016).

and

$$v(\alpha) = \begin{cases} 0 & \text{if } \alpha < 0, \\ 2\alpha & \text{if } \alpha \geq 0, \end{cases}$$

which is not differentiable at $\alpha = 0$.

The Lagrange function is given by

$$L(x, y, \alpha) = f^0(x, \alpha) + x + y(-x + \alpha).$$

If $y^* \in Y^*(\alpha)$, then by the first-order condition $L_x(\alpha, y^*, \alpha) = f_x^0(\alpha, \alpha) + 1 - y^* = 0$, where $f_x^0(\alpha, \alpha) = 0$, we have $y^* = 1$, and indeed, $1 \in Y^*(\alpha)$. Therefore,

$$Y^*(\alpha) = \{1\}$$

for all α.

As discussed in Example 2.1, f_α (and thus L_α) is not jointly continuous in (x, α) at $(x, \alpha) = (0, 0)$. Thus Assumption (b) in Proposition 4.1 is violated. □

Example 4.2. Let $X = \mathbb{R}$, $A = \mathbb{R}$, and $k = 1$, and let continuous functions f and g be defined by

$$\begin{aligned} f(x, \alpha) &= x, \\ g(x, \alpha) &= g^0(x, \alpha) - x, \end{aligned}$$

where

$$g^0(x, \alpha) = \begin{cases} -\dfrac{1}{\alpha^3} x^2 (x + 2\alpha)^2 & \text{if } 0 < x < -2\alpha, \\ \dfrac{1}{\alpha^3} x^2 (x - 2\alpha)^2 & \text{if } 0 < x < 2\alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $g^0(x, \alpha) = f^0(x, \alpha)$ if $0 < x < 2\alpha$ and $g^0(x, \alpha) = f^0(x, -\alpha)$ if $0 < x < -2\alpha$, where f^0 is the function defined in Example 2.1, and so refer to Figure 1(2) for the shape of the function g^0 for these ranges. One can verify that $X(\alpha) = (-\infty, 0] \cup \left[\frac{3-\sqrt{5}}{2}|\alpha|, |\alpha|\right]$, so that

$$X^*(\alpha) = \{|\alpha|\},$$

and

$$v(\alpha) = |\alpha|$$

which is not differentiable at $\alpha = 0$. As in Example 4.1, one can also verify that

$$Y^*(\alpha) = \{1\}$$

for all α (note that $g_x^0(|\alpha|, \alpha) = 0$).

A similar argument as in Example 2.1 shows that g_α (and thus L_α) is not jointly continuous in (x, α) at $(x, \alpha) = (0, 0)$. Thus Assumption (b) in Proposition 4.1 is violated. □

As a final remark, we observe that if the optimal value function v is concave, which is the case when the objective function f and the constraint functions g_i are concave jointly in x and α, then the continuity of f_α and g_α is not needed and the existence of a continuous selection of the Kuhn-Tucker vectors is sufficient to guarantee the differentiability of v. This result extends Corollary 3 in Milgrom and Segal (2002) to the case of parametric constraints.[6] Here, we let X be any nonempty set, and $A \subset \mathbb{R}$ a nonempty open interval.

Proposition 4.3. *Assume that*

(a) $X^*(\alpha) \neq \emptyset$ *for all* $\alpha \in A$,

(b) Y^* *has a selection* y^* *continuous at* $\bar{\alpha}$,

(c) *for all* $\bar{x} \in X^*(\bar{\alpha})$, f *and* g *are differentiable in* α *at* $(\bar{x}, \bar{\alpha})$, *and*

(d) v *is concave.*

Then v *is differentiable at* $\bar{\alpha}$ *with* $v'(\bar{\alpha}) = L_\alpha(\bar{x}, y^*(\bar{\alpha}), \bar{\alpha})$ *for any* $\bar{x} \in X^*(\bar{\alpha})$.

The proof is given in Appendix A.4.

If X is a compact convex subset of a topological vector space, f and g are continuous, and for each $\alpha \in A$, $f(\cdot, \alpha)$ and $g_i(\cdot, \alpha)$, $i = 1, \dots, k$, are concave and the Slater condition holds, i.e., there exists $\hat{x} \in X$ such that $g_i(\hat{x}, \alpha) > 0$ for all $i = 1, \dots, k$, then X^* and Y^* are nonempty-valued and upper semi-continuous. In this case, any selection of Y^* is continuous at $\bar{\alpha}$ whenever $Y^*(\bar{\alpha})$ is a singleton, which holds true, for example, when f and g are differentiable in x (where X is a subset of a Euclidean space) and the linear independence constraint qualification holds.

Assumption (b) in Proposition 4.3 is indispensable, as seen in the following example.

Example 4.3. Let $X = \mathbb{R}$, $A = \mathbb{R}$, and $k = 1$, and consider the problem given by

$$f(x, \alpha) = -|x|, \quad g(x, \alpha) = -(x - \alpha)^2 + 1,$$

which are concave in (x, α) and differentiable in α and for which the Slater condition holds. One can verify that

$$X^*(\alpha) = \begin{cases} \{1 + \alpha\} & \text{if } \alpha < -1, \\ \{0\} & \text{if } -1 \leq \alpha \leq 1, \\ \{-1 + \alpha\} & \text{if } \alpha > 1, \end{cases}$$

and

$$v(\alpha) = \begin{cases} 1 + \alpha & \text{if } \alpha < -1, \\ 0 & \text{if } -1 \leq \alpha \leq 1, \\ 1 - \alpha & \text{if } \alpha > 1, \end{cases}$$

[6]See also Corollary 2(i) in Marimon and Werner (2016), which is obtained under additional assumptions.

which is not differentiable at $\alpha = -1, 1$, while

$$Y^*(\alpha) = \begin{cases} \left\{\frac{1}{2}\right\} & \text{if } \alpha < -1 \text{ or } \alpha > 1, \\ \left[0, \frac{1}{2}\right] & \text{if } \alpha = -1, 1, \\ \{0\} & \text{if } -1 < \alpha < 1, \end{cases}$$

any selection of which is discontinuous at $\alpha = -1, 1$. □

5 Applications

In this section, we discuss the (non-)differentiability of the optimal value function in some economic contexts, producer and consumer theories, in connection to Proposition 2.1 and Example 2.1 in Section 2.

5.1 Differentiability of the Support Function (or the Profit Function)

For a nonempty subset X of $\mathbb{R}^n$, let π_X denote the *support function* of X, the function $\pi_X \colon \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$\pi_X(p) = \sup_{x \in X} p \cdot x,$$

and let S_X denote the associated optimal solution correspondence:

$$S_X(p) = \{x \in \mathbb{R}^n \mid x \in X,\ p \cdot x = \pi_X(p)\}.$$

If X is the production set of a firm, π_X and S_X are the profit function and the supply correspondence of the firm, respectively.

For support functions, it is well known that, if X is closed and convex, differentiability holds precisely when the optimal solution is unique.

Theorem 5.1. *Assume that X is a nonempty closed convex subset of $\mathbb{R}^n$. Then π_X is differentiable at $\bar{p}$ if and only if $S_X(\bar{p})$ is a singleton, in which case $S_X(\bar{p}) = \{\nabla \pi_X(\bar{p})\}$.*

In convex analysis (for example in Rockafellar (1970, Chapter 25)), this theorem is proved via conjugate duality for convex functions. Here, we present an alternative proof, showing that it also follows from our envelope theorem, Proposition 2.1 (or its multidimensional parameter version, Proposition A.1, or Corollary A.2, in Appendix A.1). The continuity of the partial derivative of the objective function $p \cdot x$ with respect to p is obvious in this case. The main obstacle in applying our proposition is the possible unboundedness of X, but it turns out that under the closedness and convexity of X, the linearity (or more generally, quasi-concavity) of $p \cdot x$ in x, and its continuity in (x, p), the point boundedness of S_X in fact implies the neighborhood boundedness.

Lemma 5.2. *Assume that X is a nonempty closed convex subset of $\mathbb{R}^n$. If $S_X(\bar{p})$ is nonempty and bounded, then S_X is nonempty and uniformly bounded near $\bar{p}$, i.e., there exists a neighborhood $U \subset \mathbb{R}^n$ of $\bar{p}$ such that $S_X(p) \neq \emptyset$ for all $p \in U$ and $\bigcup_{p \in U} S_X(p)$ is bounded.*

The proof is given in Appendix A.5 as a slightly more general form of Lemma A.4.

Given Lemma 5.2, the "if" part of Theorem 5.1 follows readily from our envelope theorem.

Proof of Theorem 5.1. For the "if" part, given a $\bar{p}$ for which $S_X(\bar{p})$ is a singleton, let $U \subset \mathbb{R}^n$ be a neighborhood of $\bar{p}$ as given in Lemma 5.2, and let $\hat{X}$ be the closure of $\bigcup_{p \in U} S_X(p)$, which is nonempty and compact. For all $p \in U$, we have $\pi_{\hat{X}}(p) = \pi_X(p)$ and $S_{\hat{X}}(p) = S_X(p)$ (where $\pi_{\hat{X}}$ is the support function of $\hat{X}$ and $S_{\hat{X}}$ is the associated optimal solution correspondence), and $S_{\hat{X}}$ is nonempty-valued and upper semi-continuous on U. The function $f(x, p) = p \cdot x$ is differentiable in p, and $\nabla_p f(x, p) = x$ is continuous in (x, p). Thus, it follows from Corollary A.2 that $\pi_{\hat{X}}$, and hence π_X, is differentiable at $\bar{p}$ with $\nabla \pi_X(\bar{p}) = \nabla \pi_{\hat{X}}(\bar{p}) = \nabla_p f(\bar{x}, \bar{p}) = \bar{x}$, where $\bar{x}$ is the unique element of $S_{\hat{X}}(\bar{p})$, and hence of $S_X(\bar{p})$.

For the "only if" part, first, the definition immediately implies that $S_X(\bar{p}) \subset \partial \pi_X(\bar{p})$, where $\partial \pi_X(\bar{p})$ is the set of subgradients of π_X at $\bar{p}$, which, by the convexity of π_X, equals $\{\nabla \pi_X(\bar{p})\}$ if π_X is differentiable at $\bar{p}$. Then, the nonemptiness of $S_X(\bar{p})$ follows from the fact that $S_X(\bar{p}) \supset \partial \pi_X(\bar{p})$ for closed convex set X, which is proved via the separation theorem. Alternatively, the nonemptiness of $S_X(\bar{p})$ directly follows by an elementary argument (see Lemma A.5 in Appendix A.5) from the differentiability of π_X at $\bar{p}$ (under the closedness of X), and, as we demonstrated in the Introduction, with the differentiability of π_X it is straightforward to derive the envelope formula, $\nabla \pi_X(\bar{p}) = \nabla_p f(\bar{x}, \bar{p}) = \bar{x}$ for any $\bar{x} \in S_X(\bar{p})$, which implies that $S_X(\bar{p}) = \{\nabla \pi_X(\bar{p})\}$. ∎

The convexity of X can be dropped if X is compact, in which case $\operatorname{conv} X$, the convex hull of X, is closed.

Corollary 5.3. *Assume that X is a nonempty compact subset of $\mathbb{R}^n$. Then π_X is differentiable at $\bar{p}$ if and only if $S_X(\bar{p})$ is a singleton, in which case $S_X(\bar{p}) = \{\nabla \pi_X(\bar{p})\}$.*

Proof. By the convexity of $p \cdot x$ in x, if $x \in \operatorname{conv} X$, then there exists $x' \in X$ such that $p \cdot x \leq p \cdot x'$. This implies that, for all p, $\pi_X(p) = \pi_{\operatorname{conv} X}(p)$ and

$\operatorname{conv} S_X(p) \subset S_{\operatorname{conv} X}(p)$. Conversely, suppose that $\bar{x} \in S_{\operatorname{conv} X}(p)$, so that $\bar{x} = \sum_{i=1}^{k} \lambda^i x^i$ for some $x^1, \ldots, x^k \in X$ and $\lambda^1, \ldots, \lambda^k > 0$, and $p \cdot \bar{x} \geq p \cdot x$ for all $x \in \operatorname{conv} X$. Then for any $i = 1, \ldots, k$, $\lambda^i p \cdot x^i = p \cdot \bar{x} - \sum_{j \neq i} \lambda^j p \cdot x^j \geq p \cdot \bar{x} - \sum_{j \neq i} \lambda^j p \cdot \bar{x} = \lambda^i p \cdot \bar{x}$, or $p \cdot x^i \geq p \cdot \bar{x}$, so that $x^i \in S_X(p)$. This implies that $\bar{x} \in \operatorname{conv} S_X(p)$.

Now, $\operatorname{conv} X$ is nonempty and closed if X is nonempty and compact. Therefore, it follows from Theorem 5.1 that $\pi_X = \pi_{\operatorname{conv} X}$ is differentiable at $\bar{p}$ if and only if $S_{\operatorname{conv} X}(\bar{p}) = \operatorname{conv} S_X(\bar{p})$ is a singleton, which holds true if and only if $S_X(\bar{p})$ is a singleton, in which case $S_X(\bar{p}) = \{\nabla \pi_X(\bar{p})\}$. ∎

Remark 5.1. The "if" part of Theorem 5.1 follows also from Milgrom and Segal's (2002) theorem (Theorem 3.2). Let $U \subset \mathbb{R}^n$ be as in Lemma 5.2. Then, for each $i = 1, \ldots, n$, the family of functions $p_i \mapsto (\bar{p}_{-i}, p_i) \cdot x$, $x \in \hat{X}$, on the domain $\{p_i \in \mathbb{R} \mid (\bar{p}_{-i}, p_i) \in U\}$ is equidifferentiable by the compactness of $\hat{X}$ (where $\hat{X}$ is the closure of $\bigcup_{p \in U} S_X(p)$). (The other assumptions can be verified to hold as in our proof above.) Thus by Theorem 3.2, $\pi_X = \pi_{\hat{X}}$ is differentiable for each p_i at $\bar{p}$. Since π_X is convex, the existence of all the partial derivatives implies its differentiability.

5.2 (Non-)Differentiability of the Indirect Utility Function and the Expenditure Function

In this subsection, we consider the (non-)differentiability of the optimal value function in consumer theory. First, as a corollary to our envelope theorem, Proposition 2.1 (or Corollary A.2), we provide a set of sufficient conditions under which the indirect utility and the expenditure functions are differentiable when the optimal solution is unique. Second, based on Example 2.1, we present an example of a continuous utility function with strictly positive partial derivatives for which these functions fail to be differentiable even when the solution is unique. In this example, again, the partial derivatives of the utility function fail to be continuous. In what follows, we let $X = \mathbb{R}^n_+$.

5.2.1 Differentiability of the Indirect Utility Function

Let a utility function $u \colon X \to \mathbb{R}$ be given, and consider the indirect utility function v defined on $\mathbb{R}^n_{++} \times \mathbb{R}_{++}$:

$$v(p, w) = \sup_{x \in B(p,w)} u(x),$$

where $B(p, w) = \{x \in X \mid p \cdot x \leq w\}$. Let D be the Walrasian demand correspondence, i.e., $D(p, w) = \{x \in X \mid x \in B(p, w),\ u(x) = v(p, w)\}$. We obtain as a corollary to our envelope theorem the following sufficient condition for the differentiability of v.

Proposition 5.4. *Assume that*

(a) *u is locally nonsatiated and continuous,*

(b) $D(\bar{p}, \bar{w}) = \{\bar{x}\}$, *and*

(c) *for some* j *with* $\bar{x}_j > 0$ *and for some neighborhoods* X_j^0 *and* X_{-j}^0 *of* $\bar{x}_j$ *and* $\bar{x}_{-j}$ *in* $\mathbb{R}_+$ *and* $\mathbb{R}_+^{n-1}$*, respectively,* u_{x_j} *exists on* $X_j^0 \times X_{-j}^0$ *and is continuous in* x *at* $\bar{x}$.

Then v *is differentiable at* $(\bar{p}, \bar{w})$ *with*

$$v_{p_i}(\bar{p}, \bar{w}) = -\frac{u_{x_j}(\bar{x})}{\bar{p}_j}\bar{x}_i, \quad v_w(\bar{p}, \bar{w}) = \frac{u_{x_j}(\bar{x})}{\bar{p}_j}$$

for any j *satisfying the condition in* (c).

Proof. By the local nonsatiability, the inequality constraint $p \cdot x \leq w$ can be replaced with the equality constraint $p \cdot x = w$. Let $D(\bar{p}, \bar{w}) = \{\bar{x}\}$, where $\bar{p} \cdot \bar{x} = \bar{w}$. Let j, X_j^0, and X_{-j}^0 be as in Assumption (c), where $\bar{x}_j = \frac{1}{\bar{p}_j}\left(\bar{w} - \sum_{i \neq j} \bar{p}_i \bar{x}_i\right) \in X_j^0$. Since $\frac{1}{p_j}\left(w - \sum_{i \neq j} p_i x_i\right)$ is continuous in (x_{-j}, p, w), there exist a compact neighborhood $X_{-j}^1 \subset X_{-j}^0$ of $\bar{x}_{-j}$ in $\mathbb{R}_+^{n-1}$, an open neighborhood P^0 of $\bar{p}$ in $\mathbb{R}_{++}^n$, and an open neighborhood W^0 of $\bar{w}$ in $\mathbb{R}_{++}$ such that $\frac{1}{p_j}\left(w - \sum_{i \neq j} p_i x_i\right) \in X_j^0$ for all $(x_{-n}, p, w) \in X_{-j}^1 \times P^0 \times W^0$. Since D is upper semi-continuous by the continuity of u and since $D_{-j}(\bar{p}, \bar{w}) \subset X_{-j}^1$, we can take open neighborhoods $P^1 \subset P^0$ and $W^1 \subset W^0$ of $\bar{p}$ and $\bar{w}$, respectively, such that $D_{-j}(p, w) \subset X_{-j}^1$ for all $(p, w) \in P^1 \times W^1$, where $D_{-j}(p, w) = \{x_{-j} \in \mathbb{R}_+^{n-1} \mid (x_j, x_{-j}) \in D(p, w) \text{ for some } x_j \in \mathbb{R}_+\}$.

Now define the continuous function $f\colon X_{-j}^1 \times P^1 \times W^1 \to \mathbb{R}$ by

$$f(x_{-j}, p, w) = u\left(\frac{1}{p_j}\left(w - \textstyle\sum_{i \neq j} p_i x_i\right), x_{-j}\right).$$

By construction, $v(p, w) = \max_{x_{-j} \in X_{-j}^1} f(x_{-j}, p, w)$ for all $(p, w) \in P^1 \times W^1$ and $\{x_{-j} \in X_{-j}^1 \mid f(x_{-j}, \bar{p}, \bar{w}) = v(\bar{p}, \bar{w})\} = \{\bar{x}_{-j}\}$, while by Assumption (c), $\nabla_{(p,w)} f$ exists on $X_{-j}^1 \times P^1 \times W^1$ and is continuous in (x_{-j}, p, w) at $(\bar{x}_{-j}, \bar{p}, \bar{w})$. Therefore, it follows from Corollary A.2 that v is differentiable at $(\bar{p}, \bar{w})$, and

$$v_{p_i}(\bar{p}, \bar{w}) = f_{p_i}(\bar{x}_{-j}, \bar{p}, \bar{w}) = u_{x_j}(\bar{x})\frac{1}{\bar{p}_j}(-\bar{x}_i),$$
$$v_w(\bar{p}, \bar{w}) = f_w(\bar{x}_{-j}, \bar{p}, \bar{w}) = u_{x_j}(\bar{x})\frac{1}{\bar{p}_j},$$

as claimed. ▮

5.2.2 Differentiability of the Expenditure Function

Let a utility function $u\colon X \to \mathbb{R}$ be given, and let $U \subset \mathbb{R}$ be the open interval $(\inf u(X), \sup u(X))$. We assume that $U \neq \emptyset$ (i.e., u is non-constant), and consider the expenditure function e defined on $\mathbb{R}^n_{++} \times U$:

$$e(p,u) = \inf_{x \in V(u)} p \cdot x,$$

where $V(u) = \{x \in X \mid u(x) \geq u\}$ (which is nonempty for all $u \in U$). Let H be the Hicksian demand correspondence, i.e., $H(p,u) = \{x \in X \mid x \in V(u),\ p \cdot x = e(p,u)\}$.

With $\bar{u}$ fixed, the expenditure function $e(p,\bar{u})$, being the support function of $V(\bar{u})$, is differentiable in p if and only if the solution to the expenditure minimization problem is unique, as discussed in Subsection 5.1.

Proposition 5.5. *Assume that u is continuous. Then $e(\cdot,\bar{u})$ is differentiable at $\bar{p}$ if and only if $H(\bar{p},\bar{u})$ is a singleton, in which case $H(\bar{p},\bar{u}) = \{e_p(\bar{p},\bar{u})\}$.*

Proof. Fix any $\bar{u} \in U$, and denote $\bar{e} = e(\bar{p},\bar{u})$ $(< \infty)$. Let $V^0 = V(\bar{u}) \cap \{x \in X \mid \bar{p} \cdot x \leq \bar{e} + 1\} \neq \emptyset$, which is bounded by construction and closed by the continuity of u. Take any $x^0 \in V(\bar{u})$ such that $\bar{p} \cdot x^0 \leq \bar{e} + 1/2$. By the upper semi-continuity of the correspondence $p \mapsto \{x \in X \mid p \cdot x \leq p \cdot x^0\}$, we can take an open neighborhood $U \subset \mathbb{R}^n_{++}$ of $\bar{p}$ such that $\{x \in X \mid p \cdot x \leq p \cdot x^0\} \subset \{x \in X \mid \bar{p} \cdot x \leq \bar{e} + 1\}$ for all $p \in U$. Then by construction, we have $e(p,\bar{u}) = \inf_{x \in V^0} p \cdot x$ and $H(p,\bar{u}) = \{x \in X \mid x \in V^0,\ p \cdot x = e(p,\bar{u})\}$ for all $p \in U$. Thus, by the compactness of V^0, the proposition follows from Corollary 5.3. ▮

For the differentiability of $e(p,u)$ in (p,u), we have the following sufficient condition by Proposition 5.4.

Proposition 5.6. *Assume that*

(a) *u is locally nonsatiated and continuous,*

(b) *$H(\bar{p},\bar{u}) = \{\bar{x}\}$, where $\bar{u} > u(0)$,*

(c) *for some j with $\bar{x}_j > 0$ and for some neighborhoods X^0_j and X^0_{-j} of $\bar{x}_j$ and $\bar{x}_{-j}$ in $\mathbb{R}_+$ and $\mathbb{R}^{n-1}_+$, respectively, u_{x_j} exists on $X^0_j \times X^0_{-j}$ and is continuous in x at $\bar{x}$, and*

(d) *$u_{x_j}(\bar{x}) \neq 0$ for some j satisfying the condition in* (c).

Then e is differentiable at $(\bar{p},\bar{u})$ with

$$e_{p_i}(\bar{p},\bar{u}) = \bar{x}_i, \quad e_u(\bar{p},\bar{u}) = \frac{\bar{p}_j}{u_{x_j}(\bar{x})}$$

for any j satisfying the condition in (c).

Proof. By (a), (i) e is continuous, and duality applies so that (ii) $e(p,u)$ is the solution to the equation $v(p,w)-u=0$ in w, and (iii) $D(\bar{p},\bar{w})=H(\bar{p},\bar{u})$ which equals $\{\bar{x}\}$ by (b), where $\bar{w}=e(\bar{p},\bar{u})$. Conditions (a), (iii), and (c) ensure Proposition 5.4 to hold, so that (iv) v is differentiable at $(\bar{p},\bar{w})$ with $v_{p_i}(\bar{p},\bar{w}) = -\frac{u_{x_j}(\bar{x})}{\bar{p}_j}\bar{x}_i$ and $v_w(\bar{p},\bar{w}) = \frac{u_{x_j}(\bar{x})}{\bar{p}_j}$, where j is any index that satisfies the condition in (c). By (d), we have (v) $v_w(\bar{p},\bar{w}) \neq 0$. By (i), (ii), (iv), and (v), it follows from Halkin's (1974, Theorem D) version of the implicit function theorem[7] applied to the equation $v(p,w)-u=0$ that e is differentiable at $(\bar{p},\bar{u})$ with

$$\begin{aligned} e_{p_i}(\bar{p},\bar{u}) &= -\frac{v_{p_i}(\bar{p},\bar{u})}{v_w(\bar{p},\bar{u})} = \bar{x}_i, \\ e_u(\bar{p},\bar{u}) &= -\frac{-1}{v_w(\bar{p},\bar{u})} = \frac{\bar{p}_j}{u_{x_j}(\bar{x})}, \end{aligned}$$

as claimed. ∎

5.2.3 Non-Differentiability of the Indirect Utility Function and the Expenditure Function

Now, we demonstrate that, even when the optimal solution is unique, the indirect utility function $v(p,w)$ and the expenditure function $e(p,u)$ may fail to be differentiable in wealth w and required utility u, respectively, if the continuity of the partial derivatives of the utility function is dropped in Propositions 5.4 and 5.6.

Proposition 5.7. *There exists a continuous utility function $u\colon X \to \mathbb{R}$ such that for some $\bar{p} \in \mathbb{R}^n_{++}$,*

(a) *$D(\bar{p},w)$ and $H(\bar{p},u)$ are singletons for all w and all u and are continuous (as single-valued functions) in w and u, respectively,*

(b) *u is partially differentiable in each x_i, $i=1,\dots,n$, at all $x \in X$, and*

(c) *$u_{x_i}(x) > 0$ for all $i=1,\dots,n$ and all $x \in X$,*

but $v(\bar{p},w)$ and $e(\bar{p},u)$ are not differentiable in w and u at some $\bar{w}$ and some $\bar{u}$, respectively.

Below we provide such a utility function with two commodities, which is shaped similarly to the function f given in Example 2.1 in Section 2 on the budget lines for some fixed price ratio.

Example 5.1. Let $X=\mathbb{R}^2_+$, and let the continuous function $u\colon X \to \mathbb{R}$ be defined by

$$u(x_1,x_2) = 6(x_1+x_2-2) + (x_1+x_2-2)^2 + u^0(x_1,x_2),$$

[7] See also Border (2016).

where

$$u^0(x_1,x_2) = \begin{cases} -\dfrac{8(x_1-1)^2(x_2-1)^2}{(x_1+x_2-2)^3} & \text{if } (x_1,x_2) \in [0,1)^2, \\ \dfrac{8(x_1-1)^2(x_2-1)^2}{(x_1+x_2-2)^3} & \text{if } (x_1,x_2) \in (1,\infty)^2, \\ -(x_1-1)^2(x_2-1)^2 & \text{otherwise.} \end{cases}$$

Observe that $u^0(x_1,x_2) = f\left(x_1-1, \frac{1}{2}(x_1+x_2-2)\right)$, where f is the function in Example 2.1. The function $u(x_1,x_2)$ is differentiable in each x_i at all $x \in X$, and its partial derivatives are given by

$$u_{x_i}(x_1,x_2) = 6 + 2(x_1+x_2-2) + u^0_{x_i}(x_1,x_2),$$

where

$$u^0_{x_i}(x_1,x_2) = \begin{cases} -\dfrac{8(x_i-1)(x_j-1)^2(-x_i+2x_j-1)}{(x_1+x_2-2)^4} & \text{if } (x_1,x_2) \in [0,1)^2, \\ \dfrac{8(x_i-1)(x_j-1)^2(-x_i+2x_j-1)}{(x_1+x_2-2)^4} & \text{if } (x_1,x_2) \in (1,\infty)^2, \\ -2(x_i-1)(x_j-1)^2 & \text{otherwise.} \end{cases}$$

In Supplementary Appendix B.1, we show that $u_{x_i}(x_1,x_2) > 0$, $i = 1,2$, for all $(x_1,x_2) \in X$. Thus, our utility function u is strictly increasing, so that the optimal consumption lies on the budget line. Note also that $U = (\inf u(X), \sup u(X)) = (u(0,0),\infty) = (-7,\infty)$.

In the following, we fix the price vector to $\bar{p} = (\frac{1}{2},\frac{1}{2})$. For utility maximization, substitute the budget equality $\frac{1}{2}x_1 + \frac{1}{2}x_2 = w$ into the utility function. Then, with the function f defined in Example 2.1, we have $u^0(x_1, 2w-x_1) = f(x_1-1, w-1)$ for all (x_1,w) such that $0 \le x_1 \le 2w$, and therefore,

$$v(\bar{p},w) = \max_{0 \le x_1 \le 2w} 12(w-1) + 4(w-1)^2 + f(x_1-1, w-1).$$

Thus, as in Example 2.1, for all w, the solution is $x_1^* = w$, so that the optimal consumption is given by $D(\bar{p},w) = \{(w,w)\}$, and the indirect utility function with $p = \bar{p}$ is given by

$$v(\bar{p},w) = 12(w-1) + 4(w-1)^2 + |w-1|,$$

which is not differentiable in w at $w = 1$.

For expenditure minimization, duality applies by the continuity and strict monotonicity of u. Therefore, the expenditure function for $u \in U$

with $p = \bar{p}$ is given by the inverse function of $w \mapsto v(\bar{p}, w)$: in a closed form,

$$e(\bar{p}, u) = \begin{cases} \dfrac{-3 + \sqrt{16u + 121}}{8} & \text{if } -7 < u < 0, \\ \dfrac{-5 + \sqrt{16u + 169}}{8} & \text{if } u \geq 0, \end{cases}$$

which is not differentiable in u at $u = 0$. The Hicksian demand is given by $H(\bar{p}, u) = D(\bar{p}, e(\bar{p}, u)) = \{(e(\bar{p}, u), e(\bar{p}, u))\}$.

While continuous separately in x_1 and in x_2, the partial derivatives u_{x_i} are not continuous jointly in (x_1, x_2) at $(x_1, x_2) = (1, 1)$. To see this for $i = 1$, set $x_2 = c(x_1 - 1) + 1$, and let $x_1 \to 1$. Then, if $c > 0$, then

$$\lim_{x_1 \to 1-0} u_{x_1}(x_1, c(x_1 - 1) + 1) = -\frac{8c^2(2c - 1)}{(c + 1)^4},$$
$$\lim_{x_1 \to 1+0} u_{x_1}(x_1, c(x_1 - 1) + 1) = \frac{8c^2(2c - 1)}{(c + 1)^4},$$

while $\lim_{x_1 \to 1} u_{x_1}(x_1, c(x_1 - 1) + 1) = 0$ otherwise (and $\lim_{x_2 \to 1} u_{x_1}(1, x_2) = 0$). Thus, Assumption (c) in Propositions 5.4 and 5.6 is violated. ∎

Remark 5.2. Clearly, this example can also be read as an example of a failure of the differentiability of the cost function in production. With the function u above, define the production function by $F(L, K) = u(L, K) - u(0, 0)$ (where $u(0, 0) = -7$). With the factor price vector fixed to $(\bar{w}, \bar{r}) = (\frac{1}{2}, \frac{1}{2})$, the cost function for this production function is given by $c(y) = e((\bar{w}, \bar{r}), y - 7)$, which is not differentiable at $y = 7$. Furthermore, this example also demonstrates that the property that the short-run cost curves are tangent to the long-run cost curve (in the strict sense that both are differentiable and share a common tangent line) may fail even when the short-run cost function is differentiable in y and the cost minimizer is unique. Note that, in this example, since with K fixed, $F(L, K)$ is continuously differentiable in L and $F_L(L, K) > 0$, the inverse function $y \mapsto L(y; K)$ of $L \mapsto F(L, K)$ is well defined and continuously differentiable, and thus the short-run cost function $c(y; K) = \bar{w}L(y; K) + \bar{r}K$ is differentiable in y.

Remark 5.3. The function u in the above example is not (totally) differentiable in x at $x = (1, 1)$. One can in fact modify the example to be differentiable in x everywhere, thereby strengthening Proposition 5.7 by replacing Condition (b) with the condition "u is differentiable in x at all $x \in X$". In Supplementary Appendix B.2, we present such a function, as well as the corresponding totally differentiable version of Example 2.1.

Appendix

A.1 Proof of Proposition 2.1

In this section, we prove a slight generalization of Proposition 2.1 where α is a multidimensional variable, which is used in Section 5. Let A be a nonempty open subset of $\mathbb{R}^m$ (instead of one of $\mathbb{R}$). The function $v\colon A \to (-\infty, \infty]$ is differentiable at $\alpha \in A$ if $v(\alpha) < \infty$, and there exists $q \in \mathbb{R}^m$ such that for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\frac{|v(\alpha + z) - v(\alpha) - q \cdot z|}{|z|} \leq \varepsilon$$

whenever $|z| \leq \delta$, $z \neq 0$. In this case, q is called the derivative, or gradient, of v at α and is denoted $\nabla v(\alpha)$, which equals $(v_{\alpha_1}(\alpha), \dots, v_{\alpha_m}(\alpha))'$. The gradient of the function $\alpha \mapsto f(x, \alpha)$ at (x, α), if it exists, will be denoted $\nabla_\alpha f(x, \alpha)$.

Proposition A.1. *Assume that*

(a) *X^* has a selection x^* continuous at $\bar{\alpha}$, and*

(b) *f is differentiable in α in a neighborhood of $(x^*(\bar{\alpha}), \bar{\alpha})$, and $\nabla_\alpha f$ is continuous in (x, α) at $(x^*(\bar{\alpha}), \bar{\alpha})$.*

Then v is differentiable at $\bar{\alpha}$ with $\nabla v(\bar{\alpha}) = \nabla_\alpha f(x^*(\bar{\alpha}), \bar{\alpha})$.

Proof. Let x^* be a selection of X^* continuous at $\bar{\alpha}$ as assumed, and denote $x^*(\bar{\alpha}) = \bar{x}$. By the definition of v, for any $z \in \mathbb{R}^m$ with $\bar{\alpha} + z \in A$ we have

$$\begin{aligned} f(\bar{x}, \bar{\alpha} + z) - f(\bar{x}, \bar{\alpha}) &\leq v(\bar{\alpha} + z) - v(\bar{\alpha}) \\ &\leq f(x^*(\bar{\alpha} + z), \bar{\alpha} + z) - f(x^*(\bar{\alpha} + z), \bar{\alpha}). \end{aligned}$$

Let $\delta_0 > 0$ be such that for all z with $|z| \leq \delta_0$ and all $\theta \in [0, 1]$, $(x^*(\bar{\alpha} + z), \bar{\alpha} + \theta z)$ lies in the neighborhood of $(\bar{x}, \bar{\alpha})$ in which f is differentiable in α. By the mean value theorem, for any z with $|z| \leq \delta_0$, we have

$$f(x^*(\bar{\alpha} + z), \bar{\alpha} + z) - f(x^*(\bar{\alpha} + z), \bar{\alpha}) = \nabla_\alpha f(x^*(\bar{\alpha} + z), \bar{\alpha} + \theta z) \cdot z$$

for some $\theta \in (0, 1)$.

Now fix any $\varepsilon > 0$. By the differentiability of $f(\bar{x}, \cdot)$ at $\bar{\alpha}$, we can take a $\delta_1 > 0$ such that if $|z| \leq \delta_1$, $z \neq 0$, then

$$\frac{|f(\bar{x}, \bar{\alpha} + z) - f(\bar{x}, \bar{\alpha}) - \nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot z|}{|z|} \leq \varepsilon.$$

By the continuity of $\nabla_\alpha f$ at $(\bar{x}, \bar{\alpha})$ as well as that of x^* at $\bar{\alpha}$, we can also take a $\delta_2 > 0$, where $\delta_2 \leq \delta_0$, such that if $|z| \leq \delta_2$ and $|z'| \leq \delta_2$, then

$$|\nabla_\alpha f(x^*(\bar{\alpha} + z), \bar{\alpha} + z') - \nabla_\alpha f(\bar{x}, \bar{\alpha})| \leq \varepsilon.$$

Let $\delta = \min\{\delta_1, \delta_2\} > 0$. Then it follows that if $|z| \le \delta$, $z \ne 0$, then

$$\frac{\nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot z}{|z|} - \varepsilon \le \frac{v(\bar{\alpha} + z) - v(\bar{\alpha})}{|z|} \le \frac{\nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot z}{|z|} + \varepsilon,$$

or

$$\frac{|v(\bar{\alpha} + z) - v(\bar{\alpha}) - \nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot z|}{|z|} \le \varepsilon.$$

This means that v is differentiable at $\bar{\alpha}$, and its derivative satisfies $\nabla v(\bar{\alpha}) = \nabla_\alpha f(\bar{x}, \bar{\alpha})$. ▮

Assumption (a) in Proposition A.1 holds if X^* is nonempty-valued and upper semi-continuous and $X^*(\bar{\alpha})$ is a singleton, in which case any selection of X^* is continuous at $\bar{\alpha}$.

Corollary A.2. *Assume that*

(a) *X^* is upper semi-continuous with $X^*(\alpha) \ne \emptyset$ for all $\alpha \in A$, and $X^*(\bar{\alpha}) = \{\bar{x}\}$, and*

(b) *f is differentiable in α in a neighborhood of $(\bar{x}, \bar{\alpha})$, and $\nabla_\alpha f$ is continuous in (x, α) at $(\bar{x}, \bar{\alpha})$.*

Then v is differentiable at $\bar{\alpha}$ with $\nabla v(\bar{\alpha}) = \nabla_\alpha f(\bar{x}, \bar{\alpha})$.

A.2 Proof of Theorem 3.1

In this section, we prove the directional differentiability part of Theorem 3.1 with a multidimensional parameter α. As in Section A.1, let A be a nonempty open subset of $\mathbb{R}^m$ (instead of one of $\mathbb{R}$). For the function $v\colon A \to (-\infty, \infty]$ and for $\alpha \in A$ with $v(\alpha) < \infty$, the (one-sided) directional derivative of v at α with respect to d is defined by

$$v'(\alpha; d) = \lim_{\lambda \searrow 0} \frac{v(\alpha + \lambda d) - v(\alpha)}{\lambda}$$

if it exists (with ∞ and $-\infty$ being allowed as limits); v is directionally differentiable at α if $v(\alpha) < \infty$ and $v'(\alpha; d)$ exists and is finite for all $d \in \mathbb{R}^m$.

Theorem A.3 (Danskin)**.** *Assume that*

(a) *X^* is upper semi-continuous with $X^*(\alpha) \ne \emptyset$ for all $\alpha \in A$, and $X^*(\bar{\alpha})$ is compact, and*

(b) *for all $\bar{x} \in X^*(\bar{\alpha})$, f is differentiable in α in a neighborhood of $(\bar{x}, \bar{\alpha})$, and $\nabla_\alpha f$ is continuous in (x, α) at $(\bar{x}, \bar{\alpha})$.*

Then v is directionally differentiable at $\bar{\alpha}$ with

$$v'(\bar{\alpha}; d) = \max_{\bar{x} \in X^*(\bar{\alpha})} \nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot d.$$

Proof. Fix any $d \in \mathbb{R}^m$. By the definition of v, for all $\lambda > 0$ we have

$$\frac{f(\bar{x}, \bar{\alpha} + \lambda d) - f(\bar{x}, \bar{\alpha})}{\lambda} \leq \frac{v(\bar{\alpha} + \lambda d) - v(\bar{\alpha})}{\lambda}$$

for all $\bar{x} \in X^*(\bar{\alpha})$. By the differentiability of $f(\bar{x}, \cdot)$ at $\bar{\alpha}$, we have

$$\nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot d \leq \liminf_{\lambda \searrow 0} \frac{v(\bar{\alpha} + \lambda d) - v(\bar{\alpha})}{\lambda} \tag{5}$$

for all $\bar{x} \in X^*(\bar{\alpha})$.

Next, let $\{\lambda^k\}_{k=1}^\infty$ be a sequence such that $\lambda^k \searrow 0$ and

$$\limsup_{\lambda \searrow 0} \frac{v(\bar{\alpha} + \lambda d) - v(\bar{\alpha})}{\lambda} = \lim_{k \to \infty} \frac{v(\bar{\alpha} + \lambda^k d) - v(\bar{\alpha})}{\lambda^k}.$$

In the following, we write $\alpha^k = \bar{\alpha} + \lambda^k d$. For each k, take any $x^*(\alpha^k) \in X^*(\alpha^k)$. Then, since X^* is upper semi-continuous and $X^*(\bar{\alpha})$ is compact, $\{x^*(\alpha^k)\}_{k=1}^\infty$ has a cluster point (or limit point) in $X^*(\bar{\alpha})$ (see, e.g., Aliprantis and Border (2006, Theorem 17.16)), i.e., there exists $\bar{x} \in X^*(\bar{\alpha})$ such that for any neighborhood U of $\bar{x}$ and any K, there exists $k \geq K$ such that $x^*(\alpha^k) \in U$. Take any such $\bar{x} \in X^*(\bar{\alpha})$.

Fix any $\varepsilon > 0$. Let K_1 be such that

$$\limsup_{\lambda \searrow 0} \frac{v(\bar{\alpha} + \lambda d) - v(\bar{\alpha})}{\lambda} \leq \frac{v(\alpha^k) - v(\bar{\alpha})}{\lambda^k} + \frac{\varepsilon}{2}$$

for all $k \geq K_1$. By Assumption (b), we can take a neighborhood U of $\bar{x}$ and $\delta > 0$ such that for all $x \in U$ and all $\alpha \in A$ with $|\alpha - \bar{\alpha}| \leq \delta$, f is differentiable and f_α satisfies

$$|(\nabla_\alpha f(x, \alpha) - \nabla_\alpha f(\bar{x}, \bar{\alpha})) \cdot d| \leq \frac{\varepsilon}{2}.$$

Let K_2 be such that $|\alpha^k - \bar{\alpha}| \leq \delta$ for all $k \geq K_2$. Now take any $\ell \geq \max\{K_1, K_2\}$ such that $x^*(\alpha^\ell) \in U$. Then we have

$$\begin{aligned}
\limsup_{\lambda \searrow 0} \frac{v(\bar{\alpha} + \lambda d) - v(\bar{\alpha})}{\lambda} &\leq \frac{v(\alpha^\ell) - v(\bar{\alpha})}{\lambda^\ell} + \frac{\varepsilon}{2} \\
&\leq \frac{f(x^*(\alpha^\ell), \alpha^\ell) - f(x^*(\alpha^\ell), \bar{\alpha})}{\lambda^\ell} + \frac{\varepsilon}{2} \\
&= \nabla_\alpha f(x^*(\alpha^\ell), \bar{\alpha} + \theta(\alpha^\ell - \bar{\alpha})) \cdot d + \frac{\varepsilon}{2} \\
&\leq \nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot d + \varepsilon
\end{aligned}$$

with some $\theta \in (0, 1)$, where the second inequality follows from the definition of v, and the equality from the mean value theorem. Since $\varepsilon > 0$ has been taken arbitrarily, it follows that

$$\limsup_{\lambda \searrow 0} \frac{v(\bar{\alpha} + \lambda d) - v(\bar{\alpha})}{\lambda} \leq \nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot d, \tag{6}$$

where $\bar{x} \in X^*(\bar{\alpha})$.

By (5) and (6), $v'(\bar{\alpha}; d)$ exists and equals $\nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot d$ for some $\bar{x} \in X^*(\bar{\alpha})$. Furthermore, by (5) we have

$$v'(\bar{\alpha}; d) = \max_{\bar{x} \in X^*(\bar{\alpha})} \nabla_\alpha f(\bar{x}, \bar{\alpha}) \cdot d,$$

as desired. ▮

A.3 Proof of Proposition 3.3

Proof. Fix any $\bar{x} \in X$ and any $\varepsilon > 0$. By the equidifferentiability of $\{f(x, \cdot)\}_{x \in X}$, we can take an α^1 such that

$$\left| \frac{f(x, \alpha^1) - f(x, \bar{\alpha})}{\alpha^1 - \bar{\alpha}} - f_\alpha(x, \bar{\alpha}) \right| \leq \frac{\varepsilon}{3}$$

for all $x \in X$. By the continuity of f in x, we can take a neighborhood U of $\bar{x}$ such that for each $i = 0, 1$,

$$|f(x, \alpha^i) - f(\bar{x}, \alpha^i)| \leq \frac{\varepsilon}{6} |\alpha^1 - \bar{\alpha}|$$

for all $x \in U$.

Now let $x \in U$. Then we have

$$\begin{aligned} |f_\alpha(x, \bar{\alpha}) - f_\alpha(\bar{x}, \bar{\alpha})| \leq & \left| f_\alpha(x, \bar{\alpha}) - \frac{f(x, \alpha^1) - f(x, \bar{\alpha})}{\alpha^1 - \bar{\alpha}} \right| \\ & + \frac{|f(x, \alpha^1) - f(\bar{x}, \alpha^1)|}{|\alpha^1 - \bar{\alpha}|} + \frac{|f(\bar{x}, \bar{\alpha}) - f(x, \bar{\alpha})|}{|\alpha^1 - \bar{\alpha}|} \\ & + \left| \frac{f(\bar{x}, \alpha^1) - f(\bar{x}, \bar{\alpha})}{\alpha^1 - \bar{\alpha}} - f_\alpha(\bar{x}, \bar{\alpha}) \right| \leq \varepsilon. \end{aligned}$$

This means that $f_\alpha(\cdot, \bar{\alpha})$ is continuous. ▮

A.4 Proof of Proposition 4.3

Proof. Let x^* be a selection of X^*, and y^* a selection of Y^* continuous at $\bar{\alpha}$, and denote $x^*(\bar{\alpha}) = \bar{x}$ and $y^*(\bar{\alpha}) = \bar{y}$. Let $\bar{\varepsilon} > 0$ be such that $(\bar{\alpha} - \bar{\varepsilon}, \bar{\alpha} + \bar{\varepsilon}) \subset A$, and take any $\varepsilon \in (0, \bar{\varepsilon}]$. Then we have

$$\begin{aligned} & v(\bar{\alpha}) - v(\bar{\alpha} - \varepsilon) \\ & \quad = L(\bar{x}, \bar{y}, \bar{\alpha}) - L(\bar{x}, y^*(\bar{\alpha} - \varepsilon), \bar{\alpha}) \\ & \qquad + L(\bar{x}, y^*(\bar{\alpha} - \varepsilon), \bar{\alpha}) - L(\bar{x}, y^*(\bar{\alpha} - \varepsilon), \bar{\alpha} - \varepsilon) \\ & \qquad + L(\bar{x}, y^*(\bar{\alpha} - \varepsilon), \bar{\alpha} - \varepsilon) - L(x^*(\bar{\alpha} - \varepsilon), y^*(\bar{\alpha} - \varepsilon), \bar{\alpha} - \varepsilon), \end{aligned}$$

where the first and the third terms are nonpositive by the saddle point property, and hence we have

$$v(\bar{\alpha}) - v(\bar{\alpha} - \varepsilon) \le L(\bar{x}, y^*(\bar{\alpha} - \varepsilon), \bar{\alpha}) - L(\bar{x}, y^*(\bar{\alpha} - \varepsilon), \bar{\alpha} - \varepsilon),$$

so that

$$\begin{aligned}\frac{v(\bar{\alpha}) - v(\bar{\alpha} - \varepsilon)}{\varepsilon} \le & \frac{f(\bar{x}, \bar{\alpha}) - f(\bar{x}, \bar{\alpha} - \varepsilon)}{\varepsilon} \\ & + \sum_{i=1}^{k} y_i^*(\bar{\alpha} - \varepsilon) \frac{g_i(\bar{x}, \bar{\alpha}) - g_i(\bar{x}, \bar{\alpha} - \varepsilon)}{\varepsilon}. \quad (7)\end{aligned}$$

Similarly, we have

$$\begin{aligned} & v(\bar{\alpha} + \varepsilon) - v(\bar{\alpha}) \\ & \quad = L(x^*(\bar{\alpha} + \varepsilon), y^*(\bar{\alpha} + \varepsilon), \bar{\alpha} + \varepsilon) - L(\bar{x}, y^*(\bar{\alpha} + \varepsilon), \bar{\alpha} + \varepsilon) \\ & \qquad + L(\bar{x}, y^*(\bar{\alpha} + \varepsilon), \bar{\alpha} + \varepsilon) - L(\bar{x}, y^*(\bar{\alpha} + \varepsilon), \bar{\alpha}) \\ & \qquad + L(\bar{x}, y^*(\bar{\alpha} + \varepsilon), \bar{\alpha}) - L(\bar{x}, \bar{y}, \bar{\alpha}) \\ & \quad \ge L(\bar{x}, y^*(\bar{\alpha} + \varepsilon), \bar{\alpha} + \varepsilon) - L(\bar{x}, y^*(\bar{\alpha} + \varepsilon), \bar{\alpha}) \end{aligned}$$

by the saddle point property, so that

$$\begin{aligned}\frac{v(\bar{\alpha} + \varepsilon) - v(\bar{\alpha})}{\varepsilon} \ge & \frac{f(\bar{x}, \bar{\alpha} + \varepsilon) - f(\bar{x}, \bar{\alpha})}{\varepsilon} \\ & + \sum_{i=1}^{k} y_i^*(\bar{\alpha} + \varepsilon) \frac{g_i(\bar{x}, \bar{\alpha} + \varepsilon) - g_i(\bar{x}, \bar{\alpha})}{\varepsilon}. \quad (8)\end{aligned}$$

By the concavity of v, we also have

$$\frac{v(\bar{\alpha}) - v(\bar{\alpha} - \varepsilon)}{\varepsilon} \ge \frac{v(\bar{\alpha} + \varepsilon) - v(\bar{\alpha})}{\varepsilon}. \quad (9)$$

Therefore, from (7)–(9) we have

$$\begin{aligned} & \frac{f(\bar{x}, \bar{\alpha}) - f(\bar{x}, \bar{\alpha} - \varepsilon)}{\varepsilon} + \sum_{i=1}^{k} y_i^*(\bar{\alpha} - \varepsilon) \frac{g_i(\bar{x}, \bar{\alpha}) - g_i(\bar{x}, \bar{\alpha} - \varepsilon)}{\varepsilon} \\ & \ge \frac{v(\bar{\alpha}) - v(\bar{\alpha} - \varepsilon)}{\varepsilon} \ge \frac{v(\bar{\alpha} + \varepsilon) - v(\bar{\alpha})}{\varepsilon} \\ & \ge \frac{f(\bar{x}, \bar{\alpha} + \varepsilon) - f(\bar{x}, \bar{\alpha})}{\varepsilon} + \sum_{i=1}^{k} y_i^*(\bar{\alpha} + \varepsilon) \frac{g_i(\bar{x}, \bar{\alpha} + \varepsilon) - g_i(\bar{x}, \bar{\alpha})}{\varepsilon} \end{aligned}$$

for all $\varepsilon \in (0, \bar{\varepsilon}]$. Now let $\varepsilon \to 0$. By the differentiability of f and g in α at $(\bar{x}, \bar{\alpha})$ and the continuity of y^* at $\bar{\alpha}$, the left most and the right most terms converge to $L_\alpha(\bar{x}, \bar{y}, \bar{\alpha}) = f_\alpha(\bar{x}, \bar{\alpha}) + \sum_{i=1}^{k} \bar{y}_i g_{i\alpha}(\bar{x}, \bar{\alpha})$, and so do the two terms in between. ▮

A.5 Proofs for Section 5.1

Let $X \subset \mathbb{R}^n$ be a nonempty set and $A \subset \mathbb{R}^m$ a nonempty open set, and, as in Section 2, for a function $f\colon X \times A \to \mathbb{R}$ let X^* denote the optimal solution correspondence: $X^*(\alpha) = \{x \in \mathbb{R}^n \mid x \in X,\ f(x,\alpha) = \sup_{x' \in X} f(x',\alpha)\}$.

Lemma A.4. *Assume that*

(a) *X is closed and convex, and*

(b) *f is upper semi-continuous, $f(x,\cdot)$ is lower semi-continuous for all $x \in X$, and $f(\cdot,\alpha)$ is quasi-concave for all $\alpha \in A$.*

Then if $X^*(\bar{\alpha})$ is nonempty and bounded, then X^* is nonempty and uniformly bounded near $\bar{\alpha}$, i.e., there exists a neighborhood $U \subset A$ of $\bar{\alpha}$ such that $X^*(\alpha) \neq \emptyset$ for all $\alpha \in U$ and $\bigcup_{\alpha \in U} X^*(\alpha)$ is bounded.[8]

Proof. For $r > 0$, we denote $B_r = \{x \in \mathbb{R}^n \mid \|x\| < r\}$, $\bar{B}_r = \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$, and $S_r = \{x \in \mathbb{R}^n \mid \|x\| = r\}$.

By the boundedness of $X^*(\bar{\alpha})$, we can take an $\bar{r} > 0$ such that $X^*(\bar{\alpha}) \subset B_{\bar{r}}$. The conclusion clearly holds if $X \subset \bar{B}_{\bar{r}}$; so we assume that $X \setminus \bar{B}_{\bar{r}} \neq \emptyset$. Let $X^0 = X \cap \bar{B}_{\bar{r}}$ and $X^1 = X \cap S_{\bar{r}}$. The sets X^0 and X^1 are compact by the closedness of X, and $X^0 \neq \emptyset$ by the nonemptiness of $X^*(\bar{\alpha})$ and thus $X^1 \neq \emptyset$ by the convexity of X. Fix any $x^0 \in X^*(\bar{\alpha})$ ($\subset X^0$). By construction $X^*(\bar{\alpha}) \cap X^1 = \emptyset$, and therefore, $\max_{x \in X^1} f(x,\bar{\alpha}) - f(x^0,\bar{\alpha}) < 0$. Since $f(x,\alpha) - f(x^0,\alpha)$ is upper semi-continuous in (x,α) by (b), $\max_{x \in X^1} f(x,\alpha) - f(x^0,\alpha)$ is upper semi-continuous in α. Let $U \subset A$ be a neighborhood of $\bar{\alpha}$ such that $\max_{x \in X^1} f(x,\alpha) - f(x^0,\alpha) < 0$ for all $\alpha \in U$.

We claim that $f(x,\alpha) < f(x^0,\alpha)$ for all $x \in X \setminus X^0$ and all $\alpha \in U$. Indeed, let $x \in X \setminus X^0$ and $\alpha \in U$, and let $\lambda \in (0,1)$ be such that $x' = (1-\lambda)x^0 + \lambda x \in S_{\bar{r}}$; by the convexity of X, $x' \in X$ and hence $x' \in X^1$. Then by the construction of U, $f((1-\lambda)x^0 + \lambda x, \alpha) < f(x^0,\alpha)$, but by the quasi-concavity of $f(x,\alpha)$ in x, this implies that $f(x,\alpha) < f(x^0,\alpha)$. Thus, for all $\alpha \in U$, we have $X^*(\alpha) = \arg\max_{x \in X^0} f(x,\alpha) \subset X^0$ with $X^*(\alpha) \neq \emptyset$. ∎

Let $\pi_X(p) = \sup_{x \in X} p \cdot x$ be the support function of $X \subset \mathbb{R}^n$, and let $S_X(p) = \{x \in \mathbb{R}^n \mid x \in X,\ p \cdot x = \pi_X(p)\}$.

Lemma A.5. *Assume that X is a nonempty closed subset of $\mathbb{R}^n$. If $S_X(\bar{p})$ is empty or unbounded, then there exists $d \in \mathbb{R}^n$ such that $\pi_X(\bar{p} + \lambda d) = \infty$ for all $\lambda > 0$.*

Proof. If $S_X(\bar{p})$ is empty, let $\{x^k\}_{k=1}^\infty$ be a sequence of points in X such that $\bar{p} \cdot x^k \to \pi_X(\bar{p})$. Then $|x^k| \to \infty$; indeed, if $\{x^k\}_{k=1}^\infty$ was bounded, then its accumulation points would attain $\pi_X(\bar{p})$ on the closed set X. If $S_X(\bar{p})$ is unbounded, let $\{x^k\}_{k=1}^\infty$ simply be any sequence of points in $S_X(\bar{p})$ such

[8] A version of this lemma is also found in Hogan (1973, Theorem A.4).

that $|x^k| \to \infty$. In any case, for some coordinate $i = 1, \dots, n$ and for some subsequence of $\{x^k\}_{k=1}^{\infty}$, denoted again by $\{x^k\}_{k=1}^{\infty}$, we have $x_i^k \to \infty$ or $x_i^k \to -\infty$. Let $d = e_i$ if $x_i^k \to \infty$ or $d = -e_i$ if $x_i^k \to -\infty$, where e_i is the ith unit vector in $\mathbb{R}^n$. Then for any $\lambda > 0$, we have $(\bar{p} + \lambda d) \cdot x^k = \bar{p} \cdot x^k + \lambda |x_i^k| \to \infty$ as $k \to \infty$, and hence $\pi_X(\bar{p} + \lambda d) = \infty$. ▮

References

Aliprantis, C. D. and K. Border (2006). *Infinite Dimensional Analysis*, Third Edition, Springer-Verlag, Berlin.

Benveniste, L. M. and J. A. Scheinkman (1979). "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica* **47**, 727-732.

Border, K. C. (2015). "Miscellaneous Notes on Optimization Theory and Related Topics," `people.hss.caltech.edu/~kcb/Notes/Maximization.pdf`, v. 2016.04.11::16.04.

Border, K. C. (2016). "Notes on the Implicit Function Theorem," `people.hss.caltech.edu/~kcb/Notes/IFT.pdf`, v. 2016.01.15::00.21.

Danskin, J. M. (1966). "The Theory of Max-Min, with Applications," *SIAM Journal on Applied Mathematics* **14**, 641-664.

Danskin, J. M. (1967). *The Theory of Max-Min and Its Application to Weapons Allocation Problems*, Springer-Verlag, Berlin.

Gross, O. (1954). "The Derivatives of the Value of a Game," The RAND Corporation, Research Memorandum RM-1286.

Halkin, H. (1974). "Implicit Functions and Optimization Problems without Continuous Differentiability of the Data," *SIAM Journal on Control* **12**, 229-236.

Hogan, W. (1973). "Directional Derivatives for Extremal-Value Functions with Applications to the Completely Convex Case," *Operations Research* **21**, 188-209.

Marimon, R. and J. Werner (2016). "The Envelope Theorem, Euler and Bellman Equations, without Differentiability," mimeo.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*, Oxford University Press, Oxford.

Milgrom, P. and I. Segal (2002). "Envelope Theorems for Arbitrary Choice Sets," *Econometrica* **70**, 583-601.

Rockafellar, R. T. (1970). *Convex Analysis*, Princeton University Press, Princeton.