7. Convex Analysis

So far, we have examined a single dual problem, minimizing expenditure subject to a utility constraint. This dual problem was associated with the primal problem of maximizing utility subject to a budget constraint. Duality can be used in other settings, and this chapter provides basic tools for studying duality in concave or convex problems.

Outline

1. Properties of Convex and Concave Functions
2. Separation Theorems

7.1 Introduction to Convex Analysis

We start by examining some basic concepts of convex analysis in the simple setting of differentiable concave functions defined on the real line. Thus $f\colon \mathbb{R} \to \mathbb{R}$. The Support Property Theorem tells us that if f is differentiable at x_0, then[1]

$$f(x) \le f(x_0) + f'(x_0)(x - x_0) \tag{7.1.1}$$

By the Support Property Theorem, a differentiable function is concave if and if equation 7.1.1 holds. This important inequality is a special case of the *supergradient inequality*. It characterizes concavity for differentiable functions.

The right-hand side of equation 7.1.1 can be used to define a line,

$$y = f(x_0) + f'(x_0)(x - x_0).$$

This line has slope $f'(x_0)$ and is tangent to the graph of f at the point $\big(x_0, f(x_0)\big)$.

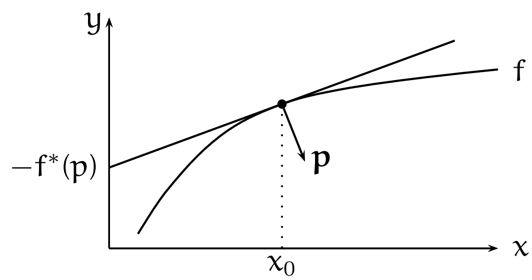

Figure 7.1.1: The tangent line at x_0 has the equation $y = f(x_0) + f'(x_0)(x - x_0)$. Because f is concave, the tangent line supports the graph of f. The graph is never above the tangent line and touches it at $\big(x_0, f(x_0)\big)$. The vector $\mathbf{p} = (p, -1)$ is perpendicular to the tangent and the vertical intercept is the concave conjugate function $f^*(p)$.

Let $p = f'(x_0)$ be the slope of the tangent line and the equation of the tangent is $p(x - x_0) = y - f(x_0)$. We now rewrite this equation in a way that that expresses it as a hyperplane in $\mathbb{R}^2$. We have

$$\mathbf{p}\cdot(x, y) = (p, -1)\cdot(x, y) = px_0 - f(x_0) = \mathbf{p}\cdot\big(x_0, f(x_0)\big). \tag{7.1.2}$$

Here $\mathbf{p} = (p, -1)$ is perpendicular to the tangent line.

[1] The full theorem, with proof is in section 31.1.1. It can also be found in Simon and Blume (1994) as Theorem 21.3.

7.1.1 The Conjugate Function

The right-hand side of equation 7.1.2 is not zero unless tangent goes through the origin. It tells us how much the tangent line is offset from the origin. That value is called the *concave conjugate function* and is denoted $f^*(p) = px_0 - f(x_0)$ when $p = f'(x_0)$.

In fact, $f^*(p)$ is the negative of the vertical intercept of the tangent line. The equation of the tangent line then becomes

$$(p, -1)\cdot(x, y) = f^*(p), \quad \text{or} \quad y = px - f^*(p)$$

and the supergradient inequality is

$$(p, -1)\cdot\big(x, f(x)\big) \geq f^*(p), \quad \text{or} \quad f(x) \leq px - f^*(p) \tag{7.1.3}$$

The left-hand side is $px - f(x)$, the vertical distance between the line and the graph of f. Since the inequality is an equality at x_0, we have

$$f^*(p) = \inf[px - f(x)].$$

The infimum is actually attained (a minimum) at x_0 when $p = f'(x_0)$.

We can immediately turn this around to ask whether a given p is the slope of a supporting tangent line. If the minimum does not exist for some value of p, it means that there is no supporting tangent line with slope p. In that case, $f^*(p) = -\infty$. If the minimum does exist, $-f^*(p)$ is the vertical intercept of the tangent line.

7.1.2 Recovering a Function from its Conjugate

Finally, if we start with a concave function f, and we know the conjugate function f^*, we know all of the supporting tangent lines via equation 7.1.3. It turns out that that is the information we need in order to find f itself. We merely take the envelope of all the tangent lines, as illustrated in Figure 7.1.2.

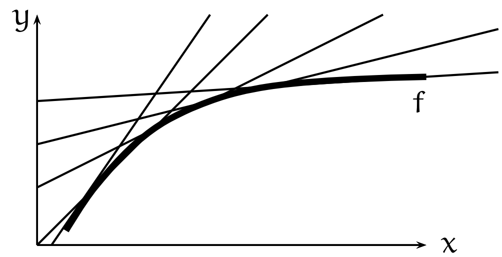

Figure 7.1.2: Since f is concave, its graph is the infimum of the tangent lines.

Convex analysis generalizes these results to the case where f is defined on $\mathbb{R}^L$ rather than $\mathbb{R}$ (this makes $\mathbf{p}$ a vector) and where f is not necessarily differentiable.

It will turn out that the envelope function is the conjugate of the conjugate—that $f^{**} = f$. We can apply these results to write expenditure, cost, and profit functions as conjugate functions, which allows us to recover the original utility and production functions as conjugates of expenditure, cost, and profit. The remainder of the chapter is devoted to that task.

7.2 Separation Theorems

Separation theorems are one of the most important mathematical tools available in economics. They play an important role in many of the key theorems of microeconomics. They pop up in duality arguments, in the welfare theorems, core equivalence, and the theory of asset pricing.

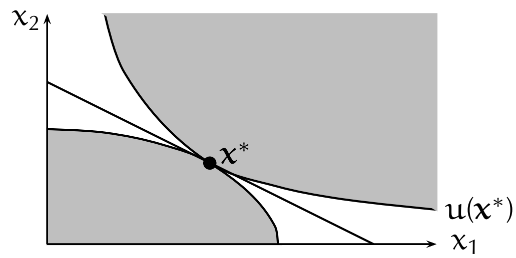

Figure 7.2.1: The optimum is at x^*. The same line is tangent to both the shaded production set and the indifference curve at x^*. We can decentralize the economy by interpreting the line as simultaneously indicating maximum profit for the producer, and as the consumer's budget constraint.

The basic idea is one that we see in intermediate micro. Suppose we have a Robinson Crusoe economy where one individual has quasiconcave preferences and a convex production possibilities set. The utility maximum is characterized by a mutual tangency of the optimal indifference curve and the production possibility set as in Figure 7.2.1. As we all know, we can use a price system to decentralize this economy. We re-interpret the tangent line as both the maximum isoprofit line for the producer and budget line for the consumer. This allows us to convert the optimal problem into an equilibrium, where the consumer maximizes utility, the firm maximizes profit, and markets clear (i.e., they choose the same point).

A separation theorem not only establishes that such a line exists in this simple case, but allows us to handle whole economies in a similar fashion.

7.2.1 Separation Theorem A

In fact, we will use four separation theorems, differing in their assumptions and the strength of the separation. The first theorem strictly separates a closed convex set from a point outside that set.

Separation Theorem A. *Suppose* $C \subset \mathbb{R}^L$ *is non-empty, closed, and convex and that* $\mathbf{x} \notin C$. *Then there is a vector* $\mathbf{p} \in \mathbb{R}^L$, $\mathbf{p} \neq \mathbf{0}$ *and a scalar* $\alpha \in \mathbb{R}$ *with* $\alpha < \mathbf{p}\cdot\mathbf{x}$ *and* $\mathbf{p}\cdot\mathbf{y} < \alpha$ *for all* $\mathbf{y} \in C$.

Proof. I claim there is a closest point in C to $\mathbf{x}$. Let $\bar{\mathbf{x}}$ be a point in C and $r = \|\bar{\mathbf{x}} - \mathbf{x}\| > 0$. Then define $B = C \cap \overline{B}_r(\mathbf{x})$. The set B is illustrated in Figure 7.2.2. As B is both closed and bounded, it is compact. By the Weierstrass Theorem, we can minimize the distance from B to $\mathbf{x}$. That also minimizes the distance from C to $\mathbf{x}$, proving the claim.

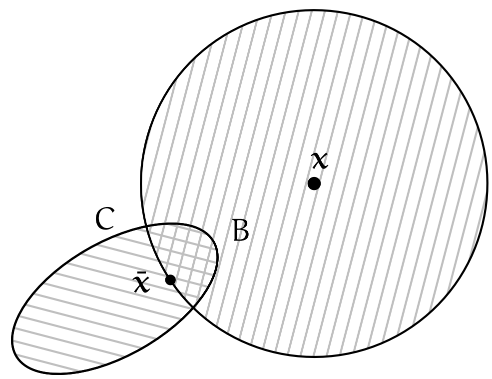

Figure 7.2.2: The set $B = C \cap \overline{B}_r(\mathbf{x})$ is cross-hatched.

7.2.2 Proof of Separation Theorem A

. Now let w be a closest point in C to x. Because C is closed, we know $w \neq x$. Define $\mathbf{p} = x - w \neq 0$.

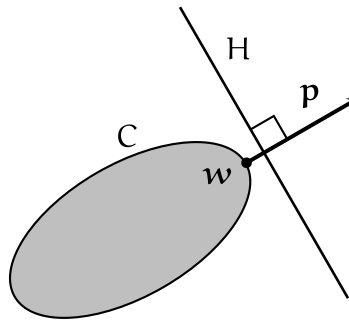

Figure 7.2.3: We now separate the convex set C from a point x by finding the nearest point in C to x (w) and using the vector $\mathbf{p}$ from w to x to perform the separation. The separating hyperplane is perpendicular to $\mathbf{p}$ and given by $H = \{z : \mathbf{p}{\cdot}z = \alpha\}$. Since C lies left and below of H, $\mathbf{p}{\cdot}y < \alpha$ for $y \in C$.

For $y \in C$, consider $\varepsilon y + (1-\varepsilon)w \in C$ for any ε with $1 > \varepsilon > 0$. It is at least as far from x as w, so

$$\|\varepsilon y + (1-\varepsilon)w - x\| \geq \|w - x\|$$

Squaring and expanding the left-hand side we obtain

$$\varepsilon^2\|y - w\|^2 + 2\varepsilon(w - x){\cdot}(y - w) + \|w - x\|^2 \geq \|w - x\|^2.$$

Cancelling $\|w - x\|^2$ from both sides and dividing by ε yields

$$\varepsilon\|y - w\|^2 + 2(w - x){\cdot}(y - w) \geq 0.$$

Let $\varepsilon \to 0$ to find $-2\mathbf{p}{\cdot}(y - w) \geq 0$. Then

$$\mathbf{p}{\cdot}w \geq \mathbf{p}{\cdot}y$$

for every $y \in C$.

Now $\mathbf{p}{\cdot}(x - w) = \|x - w\|^2 > 0$, so $\mathbf{p}{\cdot}x > \mathbf{p}{\cdot}w$. Choose any α with $\mathbf{p}{\cdot}x > \alpha > \mathbf{p}{\cdot}w$. Then

$$\mathbf{p}{\cdot}x > \alpha > \mathbf{p}{\cdot}w \geq \mathbf{p}{\cdot}y$$

for all $y \in C$, completing the proof. $\square$

7.2.3 Lemma on Interiors 1/25/22

Before proving the second separation theorem, we need some more information about convex sets. We start with a lemma showing that a strict convex combination of a point in the interior of a convex set with a point in the closure is always in the interior of the set.

Lemma 7.2.4. *Let* $C \subset \mathbb{R}^L$ *be convex. If* $x \in \operatorname{int} C$ *and* $y \in \overline{C}$*, then* $(1-\alpha)x + \alpha y \in \operatorname{int} C$ *for* $0 \le \alpha < 1$.

Proof. Let $B = \{z : \|z\| < 1\}$, so $x + \varepsilon B$ is the ball of radius ε around x. Let $0 \le \alpha < 1$. Then

$$\begin{aligned}(1-\alpha)x + \alpha y + \varepsilon B &\subset (1-\alpha)x + \alpha(C + \varepsilon B) + \varepsilon B \\ &= (1-\alpha)\left[x + \varepsilon\left(\frac{1+\alpha}{1-\alpha}\right)B\right] + \alpha C\end{aligned}$$

for all $\varepsilon > 0$. The first line uses the fact that $y \in \overline{C} \subset C + \varepsilon B$ and the second uses $\alpha < 1$. Now for ε small,

$$x + \varepsilon\left(\frac{1+\alpha}{1-\alpha}\right)B \subset C.$$

This shows $(1-\alpha)x + \alpha y + \varepsilon B \subset C$ for ε small. In other words, the ε-ball about $(1-\alpha)x + \alpha y$ is contained in C, so $(1-\alpha)x + \alpha y \in \operatorname{int} C$. $\square$

7.2.4 Interiors and Closures of Convex Sets

A corollary of this is that if a convex set has an interior, the closure of the interior is the closure of the original convex set, and the interior of the closure is the interior of the original set.

Corollary 7.2.5. *Suppose* C *is convex with* $\operatorname{int} C \neq \emptyset$*. Then* $\overline{C} = \overline{(\operatorname{int} C)}$ *and* $\operatorname{int}(\overline{C}) = \operatorname{int} C$.

Proof. Now $\operatorname{int} C \subset C$ so $\overline{(\operatorname{int} C)} \subset \overline{C}$. Let $y \in \overline{C}$ and take $x \in \operatorname{int} C$. For all $\alpha \in [0, 1)$, $(1 - \alpha)x + \alpha y \in \operatorname{int} C$ by Lemma 7.2.4. Letting $\alpha \to 1$ we see $y \in \operatorname{cl}(\operatorname{int} C)$. Thus $\overline{C} \subset \operatorname{cl}(\operatorname{int} C)$, showing $\overline{C} = \overline{(\operatorname{int} C)}$.

For the second part, $\operatorname{int} C \subset \operatorname{int}(\overline{C})$ since $C \subset \overline{C}$. Now let $y \in \operatorname{int}(\overline{C})$ and take $x \in \operatorname{int} C$. Consider $(1-\alpha)x + \alpha y$. For $\alpha > 1$ with $\alpha - 1$ small, $(1-\alpha)x + \alpha y \in \operatorname{int}(\overline{C}) \subset \overline{C}$. Setting $\beta = 1/\alpha < 1$, the lemma implies $y = (1 - \beta)x + \beta((1 - \alpha)x + \alpha y) \in \operatorname{int} C$. This shows $\operatorname{int}(\overline{C}) \subset \operatorname{int} C$ and so $\operatorname{int}(\overline{C}) = \operatorname{int} C$. $\square$

This may fail if the interior is empty or if the set is not convex.

► Example 7.2.6: Cases where Corollary 7.2.4 does not apply. In $\mathbb{R}^2$, let $C = B_1(0) \cup \{x \in \mathbb{R}^2 : x_1 \in [-2, 2],\ x_2 = 0\}$. Then $\operatorname{int} C = \operatorname{int} B_1(0)$ and $\overline{(\operatorname{int} C)} = B_1(0) \neq C = \overline{C}$.

In $\mathbb{R}^2$, let $C = \{x \in \mathbb{R}^2 : x_1 \in [-1, +1],\ x_2 = 0\}$. Then $\operatorname{int} C = \emptyset$, so $\overline{C} = C \neq \overline{(\operatorname{int} C)} = \emptyset$.

The foregoing examples show that first conclusion of Corollary 7.2.4 can fail if either of the assumptions concerning C fails.

In $\mathbb{R}$, if $C = (-1, 0) \cup (0, +1)$, $C = \operatorname{int} C$, but $\operatorname{int}(\overline{C}) = (-1, +1) \neq C$. This shows that the second conclusion of Corollary 7.2.4 can fail if C is not convex. ◄

7.2.5 Separation Theorem B

This extra information about convex sets allows us to use Separation Theorem A to prove a second separation theorem asserting that we can weakly separate two disjoint closed convex sets.

Separation Theorem B. *Suppose* $A, B \subset \mathbb{R}^L$ *are disjoint and convex. Then there is a vector* $\mathbf{p} \neq \mathbf{0}$ *and a scalar* $\alpha \in \mathbb{R}$ *with* $\mathbf{p}\cdot\mathbf{a} \leq \alpha$ *for all* $\mathbf{a} \in A$ *and* $\alpha \leq \mathbf{p}\cdot\mathbf{b}$ *for all* $\mathbf{b} \in B$.

Proof. Let $C = A - B$. Then C is convex with $\mathbf{0} \notin C$. By Corollary 7.2.4, $\mathbf{0} \notin \operatorname{int}(\overline{C}) = \operatorname{int} C \subset C$. Since $\mathbf{0} \notin \operatorname{int} C$, we can find $\mathbf{x}_n \notin \overline{C}$ with $\mathbf{x}_n \to \mathbf{0}$. Separation Theorem A yields $\mathbf{p}_n \neq \mathbf{0}$ with $0 > \mathbf{p}_n\cdot\mathbf{c}$ for all $\mathbf{c} \in C$.

Let $\mathbf{p}'_n = \mathbf{p}_n/\|\mathbf{p}_n\|$. Since the $\mathbf{p}'_n$ are bounded, then have a convergent subsequence. Let $\mathbf{p}$ be the limit. Note $\|\mathbf{p}\| = 1$. Moreover, $0 \geq \mathbf{p}\cdot\mathbf{c}$ for all $\mathbf{c} \in C$. Now $\mathbf{p}\cdot\mathbf{b} \geq \mathbf{p}\cdot\mathbf{a}$ for all $\mathbf{b} \in B$ and $\mathbf{a} \in A$. Since the $\mathbf{p}\cdot\mathbf{a}$ are bounded above, $\sup_{\mathbf{a}\in A} \mathbf{p}\cdot\mathbf{a}$ is finite. Let $\alpha = \sup \mathbf{p}\cdot\mathbf{a}$ to complete the proof. $\square$

7.2.6 Separation Theorem C

As it stands, $\mathbf{p}\cdot\mathbf{x}$ can take the value α for points in both sets. The separating hyperplane can contain points in both sets. This is not possible if one of the sets is open. We can then sharpen Separation Theorem B to get strict separation.

Separation Theorem C. *Suppose $A, B \subset \mathbb{R}^L$ are disjoint and convex and that B is open. Then there is a vector $\mathbf{p} \neq \mathbf{0}$ and a scalar $\alpha \in \mathbb{R}$ with $\mathbf{p}\cdot\mathbf{a} \leq \alpha$ for all $\mathbf{a} \in A$ and $\alpha < \mathbf{p}\cdot\mathbf{b}$ for all $\mathbf{b} \in B$.*

Proof. Separation Theorem B gives us a vector $\mathbf{p} \neq \mathbf{0}$ and $\alpha \in \mathbb{R}$ with $\mathbf{p}\cdot\mathbf{a} \leq \alpha$ for all $\mathbf{a} \in A$ and $\mathbf{p}\cdot\mathbf{b} \geq \alpha$ for all $\mathbf{b} \in B$.

Suppose there is a $\mathbf{b}_0 \in B$ with $\mathbf{p}\cdot\mathbf{b}_0 = \alpha$. Since B is open, we can find $\varepsilon > 0$ with $B_\varepsilon(\mathbf{b}_0) \subset B$. Set

$$\mathbf{b}_0 - \frac{\varepsilon}{2}\frac{\mathbf{p}}{\|\mathbf{p}\|^2} \in B_\varepsilon(\mathbf{b}_0).$$

Since this is in B,

$$\alpha \leq \mathbf{p}\cdot\left(\mathbf{b}_0 - \frac{\varepsilon}{2}\frac{\mathbf{p}}{\|\mathbf{p}\|^2}\right) = \alpha - \frac{\varepsilon}{2},$$

which is impossible. This contradiction shows that $\mathbf{p}\cdot\mathbf{b}_0 = \alpha$ is impossible, so $\mathbf{p}\cdot\mathbf{b} > \alpha$ for all $\mathbf{b} \in B$. $\square$

7.2.7 Separation of Comprehensive Convex Sets

In some cases of economic interest the separating vector has a natural sign. A set $A \subset \mathbb{R}^L$ is *comprehensive* if $\mathbf{x} \in A$ and $\mathbf{x}' \leq \mathbf{x}$ implies $\mathbf{x}' \in A$. A set $A \subset \mathbb{R}^L$ is *anti-comprehensive* if $\mathbf{x} \in A$ and $\mathbf{x}' \geq \mathbf{x}$ implies $\mathbf{x}' \in A$.

When preferences are monotonic, the upper contour set is anti-comprehensive. Free disposal will often ensure that production sets are comprehensive. If besides being convex, either A is anti-comprehensive or B is comprehensive, then the separating price vector $\mathbf{p}$ must be positive.

Corollary 7.2.7. *Suppose $A, B \subset \mathbb{R}^L$ are disjoint and convex and that either A is comprehensive or B is anti-comprehensive. If there is a vector $\mathbf{p} \neq \mathbf{0}$ and a scalar $\alpha \in \mathbb{R}$ with $\mathbf{p}\cdot\mathbf{a} \leq \alpha$ for all $\mathbf{a} \in A$ and $\alpha \leq \mathbf{p}\cdot\mathbf{b}$ for all $\mathbf{b} \in B$, then $\mathbf{p} > \mathbf{0}$.*

Proof. Suppose A is comprehensive. Take $\mathbf{a} \in A$. Then $\mathbf{a} - t\mathbf{e}_\ell \in A$ for all $t > 0$. It follows that $\mathbf{p}\cdot(\mathbf{a} - t\mathbf{e}_\ell) \geq \alpha$. Dividing by t we obtain $(\mathbf{p}\cdot\mathbf{a})/t - p_\ell \leq \alpha/t$. Let $t \to \infty$ to see that $p_\ell \geq 0$ for each $\ell = 1, \ldots, L$.

A similar argument using $\mathbf{b} + t\mathbf{e}_\ell$ works when B is anti-comprehensive. □

7.2.8 Separation Theorem D

Separation Theorem D is also an easy consequence of Separation Theorem B. It strongly separates a point not in the interior of a convex set from the interior of the set, and weakly separates the point from the entire convex set. This is often applied when the point being separated is on the boundary of the convex set.

Separation Theorem D. *Suppose* $C \subset \mathbb{R}^L$ *is non-empty and convex and that* $\mathbf{x} \notin \operatorname{int} C$. *Then there is a vector* $\mathbf{p} \neq \mathbf{0}$ *with* $\mathbf{p}\cdot\mathbf{x} < \mathbf{p}\cdot\mathbf{y}$ *for all* $\mathbf{y} \in \operatorname{int} C$ *and* $\mathbf{p}\cdot\mathbf{x} \leq \mathbf{p}\cdot\mathbf{y}$ *for all* $\mathbf{y} \in C$.

Proof. Apply Separation Theorem B to $A = C$ and $B = \{\mathbf{x}\}$ to obtain $\mathbf{p} \neq \mathbf{0}$ with $\mathbf{p}\cdot\mathbf{x} \geq \mathbf{p}\cdot\mathbf{c}$ for all $\mathbf{c} \in C$.

Now suppose $\mathbf{c} \in \operatorname{int} C$. For $\varepsilon > 0$ small enough, $\mathbf{c} + \varepsilon\mathbf{p} \in \operatorname{int} C$. Then $\mathbf{p}\cdot\mathbf{x} \geq \mathbf{p}\cdot\mathbf{c} + \varepsilon\|\mathbf{p}\|^2 > \mathbf{p}\cdot\mathbf{c}$. $\square$

7.2.9 Inverse Hicksian Demand

One application of Separation Theorem D is to the problem of inverting Hicksian demand. That is, given a consumption vector, can we find a price vector that yields that consumption as a Hicksian demand?

Separation Theorem D allows us to answer this in the affirmative.

Theorem 7.2.8. *Suppose a preference order $\succsim$ is continuous, convex, and locally non-satiated on a closed, convex consumption set $\mathfrak{X}$. If $\mathbf{x} \in \mathfrak{X}$, there is a $\mathbf{p} \neq \mathbf{0}$ with $\mathbf{p}\cdot\mathbf{x} \leq \mathbf{p}\cdot\mathbf{x}'$ for all $\mathbf{x}' \in \mathrm{R}(\mathbf{x}) = \{\mathbf{x}' : \mathbf{x}' \succsim \mathbf{x}\}$.*

Proof. By Corollary 2.4.5, the weakly preferred set $\mathrm{R}(\mathbf{x})$ is the closure of the strictly preferred set $\mathrm{P}(\mathbf{x})$. This implies $\mathrm{P}(\mathbf{x}) = \operatorname{int}\mathrm{R}(\mathbf{x})$, so $\mathbf{x} \notin \operatorname{int}\mathrm{R}(\mathbf{x})$. The set $\mathrm{R}(\mathbf{x})$ is convex by convexity of $\succsim$.

We now apply Separation Theorem D to find a $\mathbf{p} \neq \mathbf{0}$ with $\mathbf{p}\cdot\mathbf{x}' \geq \mathbf{p}\cdot\mathbf{x}$ for all $\mathbf{x}' \in \mathrm{R}(\mathbf{x})$. $\square$

When $\succsim$ is also represented by a utility function u, we can set $\bar{u} = u(\mathbf{x})$. Then Theorem 7.2.7 gives us a $\mathbf{p}$ with $\mathbf{x} \in \mathbf{h}(\mathbf{p}, \bar{u})$. Reasoning the other way, if $\bar{u}$ is given and $\bar{u}$ is in the range of u, there is a $\mathbf{p}$ with $\mathbf{x} \in \mathbf{h}(\mathbf{p}, \bar{u})$.

7.2.10 More about Inverse Hicksian Demand

There are three things to note here. First is that the price vector $\mathbf{p}$ in Theorem 7.2.7 is not unique. Any positive scalar multiple of $\mathbf{p}$ will also do the job. One way of dealing with this is to normalize prices so that $\mathbf{p}\cdot\mathbf{x} = 1$. This method is used later when examining the distance function. This method does not guarantee uniqueness, but if the utility function is differentiable and has non-zero derivative, the normalized $\mathbf{p}$ will be unique by Theorem 31.4.4. If there is a kink in the indifference curve, we cannot expect a unique price, even with normalization.

The second thing is that $\mathbf{p}$ might not be positive. Locally non-satiated preferences need not be monotonic, and if utility is decreasing in some direction at $\mathbf{x}$, the supporting price vector will not be positive. This happens when $u(x_1, x_2) = x_2 - (x_1 - 1)^2$ at the point $(3, 4)$. The supporting price vectors are positive multiples of $du(3, 4) = (-4, 1)$.

Finally, in Examples 5.2.2 and 5.2.3, we use Hicksian demands to recover the utility function from the expenditure function. Implicit in this was the idea that the Hicksian demands trace out the indifference surface, which Theorem 7.2.7 shows.

January 28, 2022