Chapter 6. Convex Sets and Their Separations

In the previous chapters, we have learned first-order necessary conditions for constrained maximization problems. We also mentioned that those conditions may not be sufficient. In this and the following two chapters, we will discuss sufficient conditions.

6.A. The Separation Property

In this chapter, we will develope a geometric approach to constrained maximization problem based on the separation property. Before digging into the details, we will illustrate the idea using an example. Consider the following maximization problem:

$$\max_x F(x)$$
$$\text{s.t. } G(x) \leq c,$$

where $G(x) \leq c$ is a scalar constraint. Let x^* denote the optimal choice, and v^* denote the maximum value. We are now interested to know the properties of the functions F and G that ensure the maximum.

Figure 6.1 illustrate the problem with two variables.

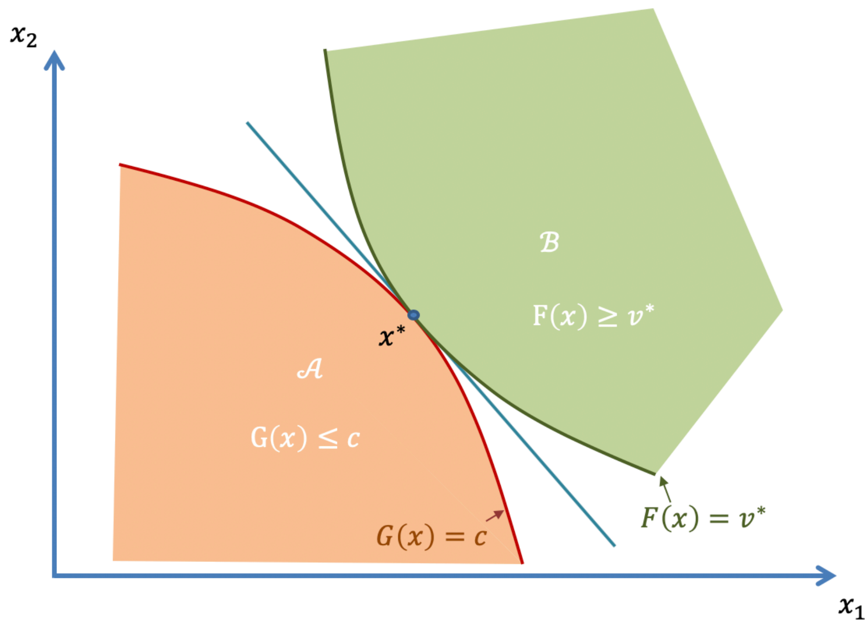


Figure 6.1: Separation by the common tangent

We know from the previous chapters that the solution is attained at the tangency point x^*. To get some idea about the general property, we will interpret the solution in terms of the curvatures of F and G. The following new concepts are needed for our discussion:

Definition 6.A.1 (Lower Contour Set)**.** For a function $f : \mathcal{S} \subset \mathbb{R}^N \to \mathbb{R}$, the *lower contour set* of f for the value $c \in \mathbb{R}$ is $\{x|f(x) \leq c\}$.

That is, the *lower contour set* is the set of all points x where $f(x) \leq c$. Similarly, we define the *upper contour set*.

Definition 6.A.2 (Upper Contour Set)**.** For a function $f : \mathcal{S} \subset \mathbb{R}^N \to \mathbb{R}$, the *upper contour set* of f for the value $c \in \mathbb{R}$ is $\{x|f(x) \geq c\}$.

Figure 6.2 provides an illustration for the one-variable case.

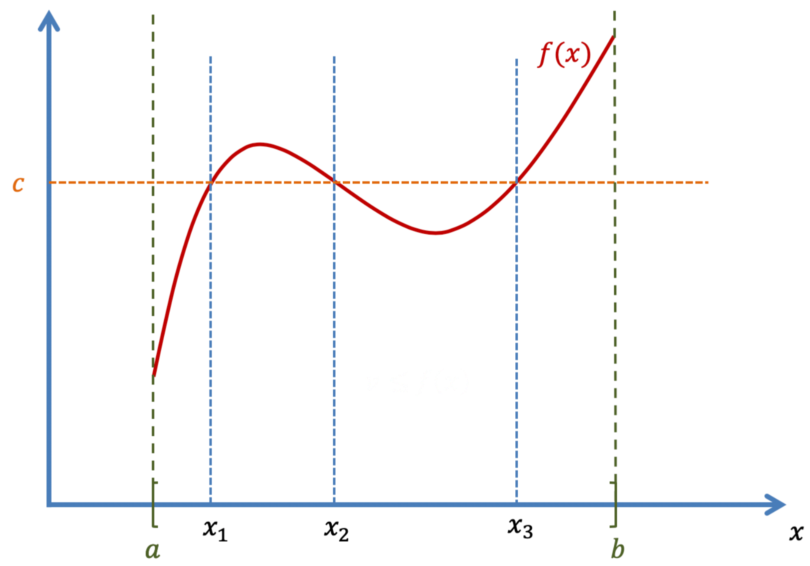


Figure 6.2: Contour Sets

- The *lower contour set* of f for the value c is $[a, x_1] \bigcup [x_2, x_3]$;
- The *upper contour set* of f for the value c is $[x_1, x_2] \bigcup [x_3, b]$.

In Figure 6.1, the green curve is $F(x) = v^*$, together with the green area, is the *upper contour set* of F for v^* (Set $\mathcal{B}$). For the constraint, the red curve $G(x) = c$, together with the orange area, is the inequality constraint $G(x) \leq c$, or the *lower contour set* of G for c (Set $\mathcal{A}$). The curvatures in Figure 6.1 ensure a maximum.

The question is, what is the general property of such curvatures? The sets $\mathcal{B}$ and $\mathcal{A}$ lie one to each side of their common tangent, with only their common point x^* on that line. In other words, the common tangent *separates* the x-plane into two halves, each containing one of the two sets. For three-variables, the common tangent is a plane. See Figure 6.3. In higher dimensions, the common tangent will be a hyperplane.

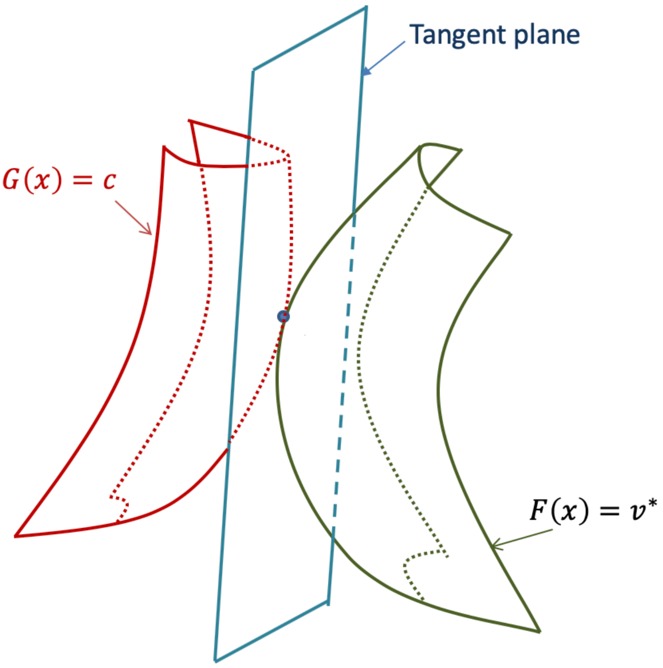


Figure 6.3: Three-variable Case

This separation property is the crucial property that allows us to find the maxima, and obtain sufficient conditions for the maximization problem. We will next examine the explicit conditions on the functions F and G that ensure the right curvature.

6.B. Convex Sets and Functions

Each of the contour sets in Figure 6.1 bulges outward, so that each bends away from the comment tangent at x^* and cannot bend back to meet the other set once again. This property is called *convexity*. Formally, Definition 6.B.1 defines *convex sets*.

Definition 6.B.1 (Convex Set)**.** A set $\mathcal{S}$ of points in n-dimensional space is called *convex* if, given any two points $x^a = (x_1^a, x_2^a, ..., x_n^a)$ and $x^b = (x_1^b, x_2^b, ..., x_n^b)$ in $\mathcal{S}$ and any real number $\alpha \in [0, 1]$, the point $\alpha x^a + (1-\alpha)x^b = (\alpha x_1^a + (1-\alpha)x_1^b, ..., \alpha x_n^a + (1-\alpha)x_n^b)$ is also in $\mathcal{S}$.

A geometric test of convexity is that given any two points of the set, the whole line segment joining them should lie in the set.

Figure 6.4 and 6.5 are examples of *convex* sets. Please be aware that to apply the geometric test of convexity, we need to ensure that for *any* two points of the set, the whole line segment lie in the set.

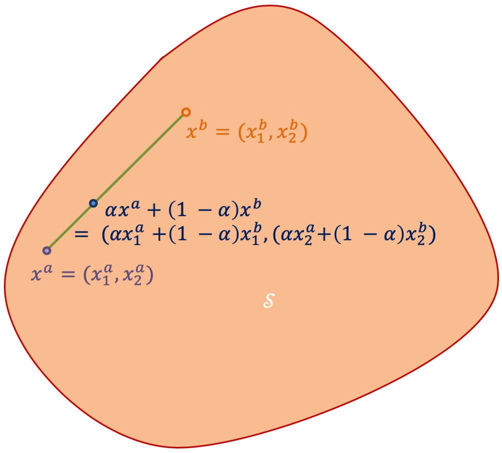


Figure 6.4: Convex Set (a)

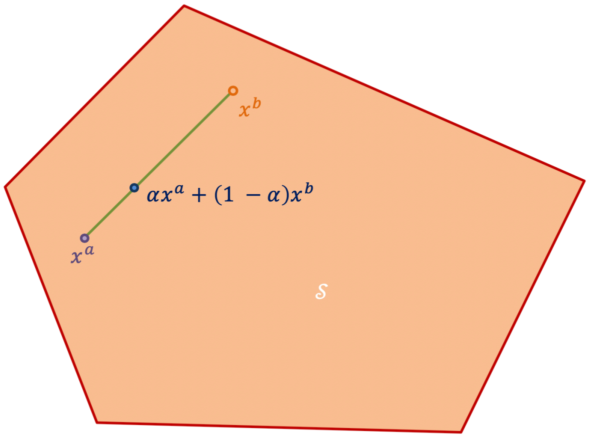


Figure 6.5: Convex Set (b)

Figure 6.6 and 6.7 are examples of *non-convex* sets. The sets are *non-convex*, since there exist points x^a and x^b and a real number α, such that the point $\alpha x^a + (1-\alpha)x^b$ is not inside the set.

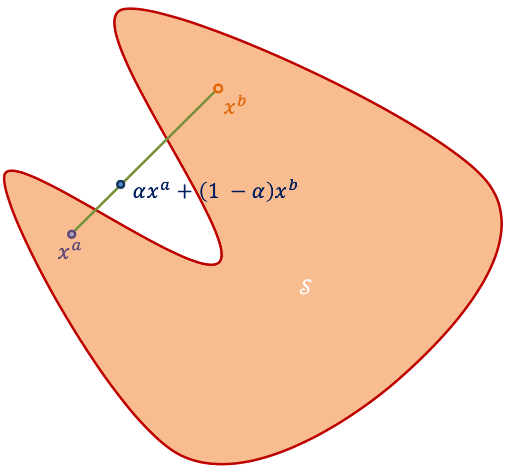


Figure 6.6: Non-Convex Set (a)

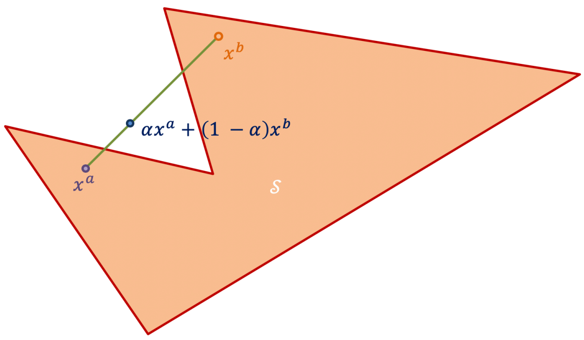


Figure 6.7: Non-Convex Set (b)

Apply the concept of *convex sets* to the *lower contour set* of G, we could reinterpret the bulging outward curvature as follows: the *lower contour set* of G is convex, or

$$\text{the set } \{x|G(x) \leq c\} \text{ is convex.} \tag{6.1}$$

This means that if x^a and x^b satisfy the constraint, so does $\alpha x^a + (1-\alpha)x^b$. Algebraically, the condition states that for all x^a and x^b that satisfies $G(x^a) \leq c$ and $G(x^b) \leq c$, and any real number $\alpha \in [0,1]$, we have $G(\alpha x^a + (1-\alpha)x^b) \leq c$. In practice, we will have to solve the maximization problem for a general value of c, so we need to invoke the condition for all c. The condition (6.1) with a general c is equivalent to

$$G(\alpha x^a + (1-\alpha)x^b) \leq \max\{G(x^a), G(x^b)\}, \tag{6.2}$$

for all x^a, x^b and for all $\alpha \in [0,1]$. A function G satisfying this condition is called *quasi-convex*. Since *quasi-convexity* is a new and important concept, we will formally state it and then prove the equivalence.

Definition 6.B.2 (Quasi-convex Function)**.** A function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is *quasi-convex* if the set $\{x|f(x) \leq c\}$ is convex for all $c \in \mathbb{R}$, or equivalently, if

$$f(\alpha x^a + (1-\alpha)x^b) \leq \max\{f(x^a), f(x^b)\}, \tag{6.3}$$

for all x^a, x^b and for all $\alpha \in [0,1]$.

Next, we show the equivalence of

(a) The set $\{x|f(x) \leq c\}$ is convex for all $c \in \mathbb{R}$;

(b) $f(\alpha x^a + (1-\alpha)x^b) \leq \max\{f(x^a), f(x^b)\}$, for all x^a, x^b and for all $\alpha \in [0,1]$.

Proof. (a) $\implies$ (b): Since (a) holds for all $c \in \mathbb{R}$, for any x^a and x^b, we could set $c = \max\{f(x^a), f(x^b)\}$. Then since $f(x^a) \leq \max\{f(x^a), f(x^b)\} = c$, $f(x^b) \leq \max\{f(x^a), f(x^b)\} = c$, by (a), we have $f(\alpha x^a + (1-\alpha)x^b) \leq c = \max\{f(x^a), f(x^b)\}$ for any $\alpha \in [0,1]$. Thus, (b) holds.

(b) $\implies$ (a): Equivalently, we show "not (a) $\implies$ not (b)".

If (a) fails, then there exists x^a, x^b, c and $\alpha \in [0,1]$ such that $f(x^a) \leq c$ and $f(x^b) \leq c$ but $f(\alpha x^a + (1-\alpha)x^b) > c$. Then $f(\alpha x^a + (1-\alpha)x^b) > c \geq \max\{f(x^a), f(x^b)\}$. Thus, (b) fails for these values of x^a, x^b and α. □

The parallel condition on F is that the *upper contour set* of F is convex, or F is *quasi-concave.* The formal definition of *quasi-concave* is given in Definition 6.B.3 below.

Definition 6.B.3 (Quasi-concave Function)**.** A function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, *quasi-concave* if the set $\{x|f(x) \geq c\}$ is convex for all $c \in \mathbb{R}$, or equivalently, if $f(\alpha x^a + (1-\alpha)x^b) \geq \min\{f(x^a), f(x^b)\}$, for all x^a, x^b and for all $\alpha \in [0,1]$.

A digression: *quasi-convexity* (*quasi-concavity*) and *convexity* (*concavity*)

The *quasi* in Definition 6.B.2 and 6.B.3 serves to distringuish them from stronger properties of *convexity* and *concavity.* Formally, we define convexity as follows.

Definition 6.B.4 (Convex Function)**.** A function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is *convex* if

$$f(\alpha x^a + (1-\alpha)x^b) \leq \alpha f(x^a) + (1-\alpha) f(x^b), \tag{6.4}$$

for all x^a, x^b and for all $\alpha \in [0,1]$.

(6.4) *convexity* implies (6.3) *quasi-convexity* since

$$\begin{aligned} f(\alpha x^a + (1-\alpha)x^b) &\underbrace{\leq}_{(6.4)} \alpha f(x^a) + (1-\alpha) f(x^b) \\ &\leq \alpha \max\{f(x^a), f(x^b)\} + (1-\alpha)\max\{f(x^a), f(x^b)\} \\ &= \max\{f(x^a), f(x^b)\}. \end{aligned}$$

In other words, a convex function must be quasi-convex.

Similarly, we could define concavity and compare it with *quasi-concavity.*

Definition 6.B.5 (Concave Function)**.** A function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is *concave* if

$$f(\alpha x^a + (1-\alpha)x^b) \geq \alpha f(x^a) + (1-\alpha) f(x^b), \tag{6.5}$$

for all x^a, x^b and for all $\alpha \in [0,1]$.

Following the same logic, we could show that a concave function must be quasi-concave.

Figure 6.8 provides a graphical illustration of a *concave* function for the one-variable case. The red dot (LHS of (6.5)) is always higher than the green dot (RHS of (6.5)).

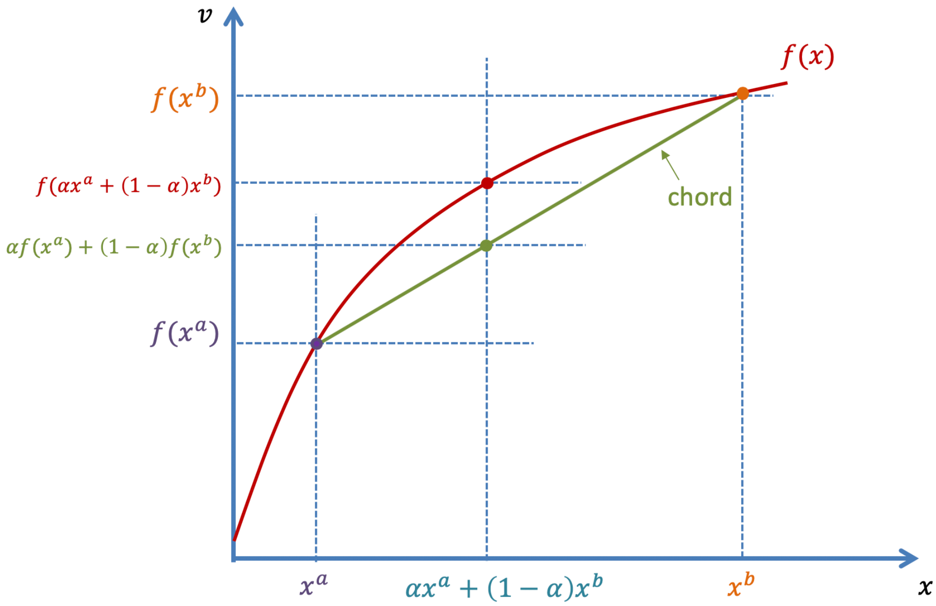


Figure 6.8: Concave Function

In words, the graph of the function lies on or above the chord joining any two points of it.

An alternative interpretation of a concave function is sometimes useful. Consider the $(n+1)$-dimensional space consisting of points like (x, v) where x is an n-dimensional vector and v is a scalar. Define the set $\mathcal{F} = \{(x, v) | v \leq f(x)\}$.

Then, we make the following claim:

Claim. f is a concave function if and only if $\mathcal{F}$ is a convex set.

Proof. " $\implies$ ": To prove that $\mathcal{F}$ is a convex set, we need to show that for all (x^a, v^a) and (x^b, v^b) that satisfy $v^a \leq f(x^a)$ and $v^b \leq f(x^b)$ and any real number $\alpha \in [0, 1]$, we have $\alpha v^a + (1-\alpha)v^b \leq f(\alpha x^a + (1-\alpha)x^b)$.

By concavity of f, we know that for all x^a and x^b and for all $\alpha \in [0, 1]$, (6.5) holds.

Therefore, for all (x^a, v^a) and (x^b, v^b) that satisfy $v^a \leq f(x^a)$ and $v^b \leq f(x^b)$ and any real number $\alpha \in [0, 1]$,

$$
\begin{aligned}
&\alpha v^a + (1-\alpha)v^b - \left[f(\alpha x^a + (1-\alpha)x^b)\right] \\
\underbrace{\leq}_{(6.5)}\ &\alpha v^a + (1-\alpha)v^b - \left[\alpha f(x^a) + (1-\alpha)f(x^b)\right] \\
\underbrace{\leq}_{v^a \leq f(x^a) \text{ and } v^b \leq f(x^b)}\ &\alpha v^a + (1-\alpha)v^b - \left[\alpha v^a + (1-\alpha)v^b\right] = 0
\end{aligned}
$$

Therefore, $\alpha v^a + (1-\alpha)v^b \leq f(\alpha x^a + (1-\alpha)x^b)$ and convexity of set $\mathcal{F}$ follows.

"$\Longleftarrow$": To prove that F is concave, we need to show that for all x^a, x^b and all $\alpha \in [0,1]$, (6.5) holds.

For any x^a and x^b, set $v^a = f(x^a)$ and $v^b = f(x^b)$, so that $v^a \leq f(x^a)$ and $v^b \leq f(x^b)$ are satisfied, i.e., $(x^a, v^a) \in \mathcal{F}$ and $(x^b, v^b) \in \mathcal{F}$. Then by convexity of set $\mathcal{F}$, for any real number $\alpha \in [0,1]$, we have $\alpha v^a + (1-\alpha)v^b \leq f(\alpha x^a + (1-\alpha)x^b) \underbrace{\Longrightarrow}_{v^a=f(x^a),\, v^b=f(x^b)} \alpha f(x^a) + (1-\alpha)f(x^b)) \leq f(\alpha x^a + (1-\alpha)x^b)$, and concavity of the function f follows. $\square$

The claim could be more easily understood graphically. Figure 6.9 illustrates the case with a scalar variable x. The function f is the red curve. The set $\mathcal{F}$ is the area shaded in orange. The claim means that the concave function f traps a convex set $\mathcal{F}$ underneath its graph. And it is clear from Figure 6.9.

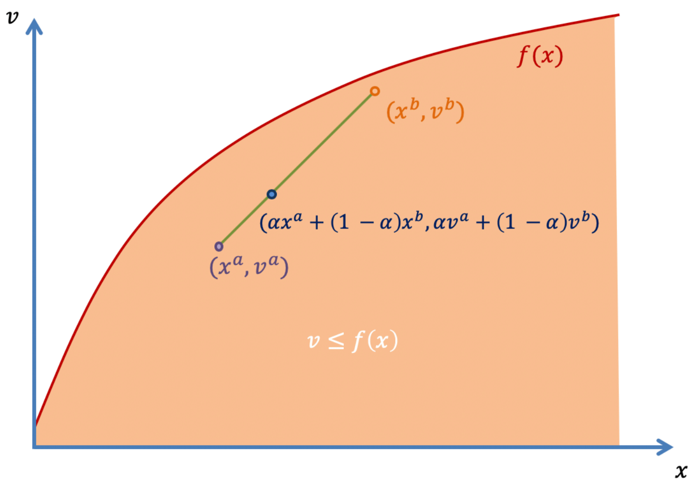


Figure 6.9: Concave Function

Two more concepts. We need to introduce two more concept before proceeding with our main result of separation.

Definition 6.B.6 (Interior Point)**.** A point $x^o \in \mathcal{S}$ is called an *interior* point if there exists a real number $r > 0$ such that for all x such that $\|x - x^o\| < r$, we have $x \in \mathcal{S}$.

That is, a point $x^o \in \mathcal{S}$ is an interior point if all points within the distance of r from the point x^o are in $\mathcal{S}$. In the plane, such points will form a disc of radius r centered at x^o.

Definition 6.B.7 (Boundary Point)**.** A point $x^o \in \mathcal{S}$ is called a *boundary* point if for any real number $r > 0$, $\exists\, x, y$ such that $\|x - x^o\| < r$, $\|y - x^o\| < r$ and $x \in \mathcal{S}$, $y \notin \mathcal{S}$.

That is, a *boundary* point of $\mathcal{S}$ is interior neither to $\mathcal{S}$ nor to the rest of the space. Figure 6.10 illustrates the interior and boundary points in a plane. In the figure,

- x^a is an *ineterior point* of $\mathcal{S}$ since we could find an open ball with radius r such that all points in the open ball are in $\mathcal{S}$.
- x^b is a *boundary point* of $\mathcal{S}$ since for any open ball around x^b, there exists points inside the ball that are in $\mathcal{S}$, and there also exists points that are outside of $\mathcal{S}$.
- x^c is neither an *ineterior point* or a *boundary point* of $\mathcal{S}$ since we could find an open ball with radius r such that all points in the open ball are outside of $\mathcal{S}$.

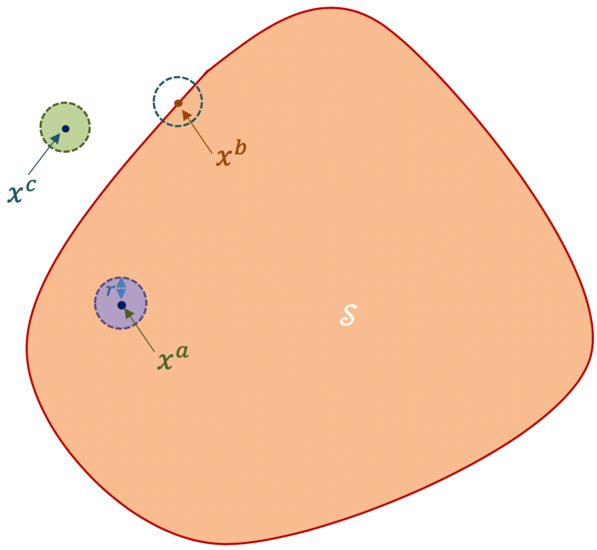


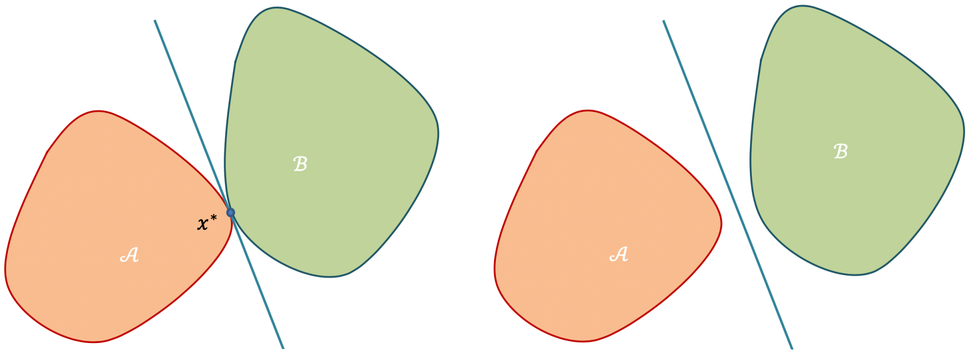
Figure 6.10: Interior and Boundary Points

For two convex sets,

(a) If the two sets only have boundary point x^* in common, the common tangent could separate them. See Figure 6.11a.

(b) If the two sets have no points in common, then there would be a clear gap between them, and we could draw a line to separate them. The line does not need to be tangent to either set. See Figure 6.11b.

(c) If the two sets have interior points in common, then we could not separate them. See Figure 6.12.



(a) common tangent (b) no points in common

Figure 6.11: Separation is possible.

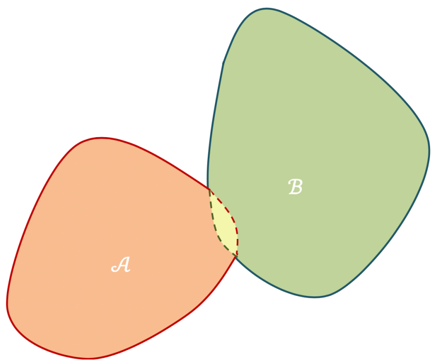
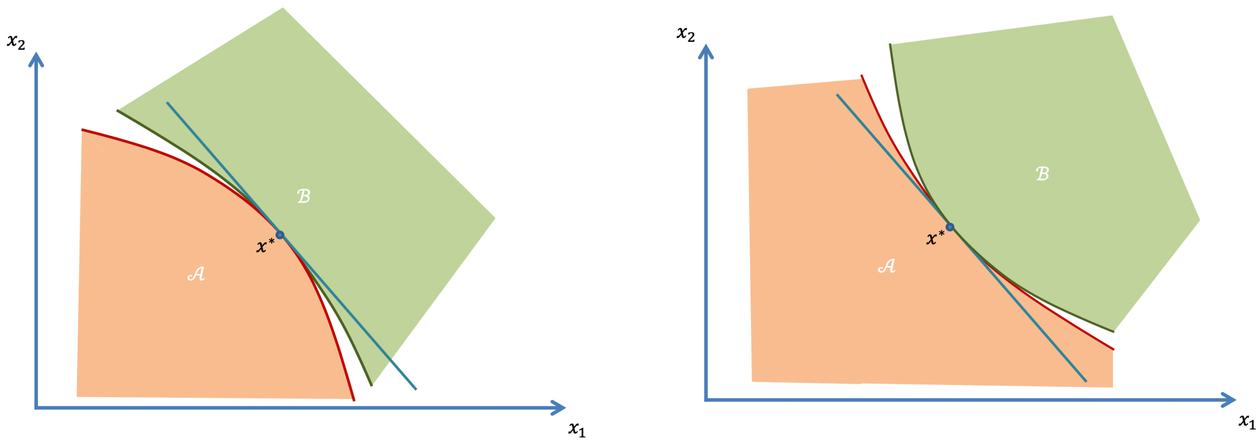


Figure 6.12: Separation is not possible.

Convexity of the sets is important. Figure 6.13 shows two cases, in each of which one set is not convex. The conmmon tangent cuts into the non-convex set, and separation fails.



(a) (b)

Figure 6.13: Partial Failure of Decentralization

The intuition from these graphs carry over to higher dimensions. We formalize the idea of separation in Theorem 6.1 below.

Theorem 6.1 (Separation Theorem)**.** *If $\mathcal{B}$ and $\mathcal{A}$ are two convex sets, that have no interior points in common, and at least one of the sets has a non-empty interior,*[1] *then we can find a non-zero vector p and a number b such that the hyperplane $px = b$ separates the two sets, or*

$$px = \sum_{i=1}^{n} p_i x_i \begin{cases} \leq b & \text{for all } x \in \mathcal{A} \\ \geq b & \text{for all } x \in \mathcal{B}. \end{cases} \tag{6.6}$$

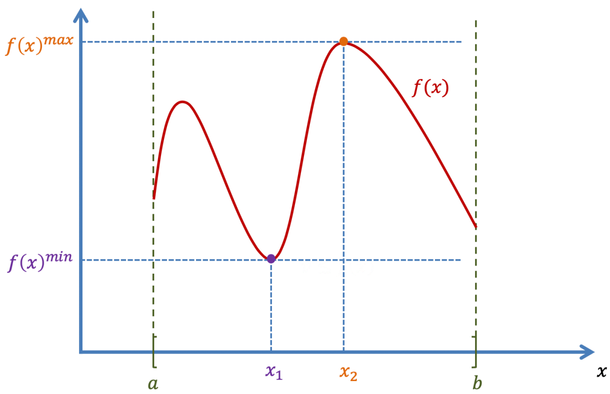
The proof of this theorem is out of the scope of this course. For interested students, a formal proof is provided in Appendix A for your reference.

6.C. Optimization by Separation

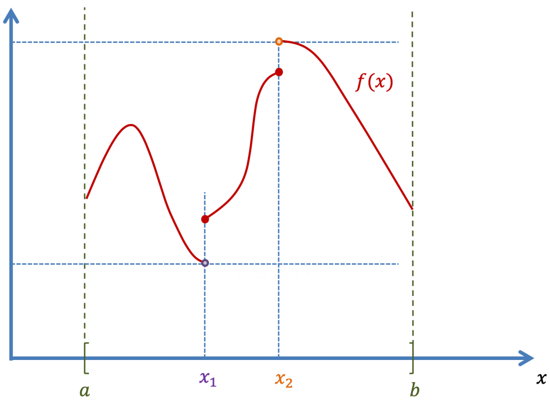
Existence of Solution. In most economic applications, the functions F and G are well-behaved, and the existence of solution is ensured by the *Extreme Value Theorem.*

Theorem (Extreme Value Theorem)**.** *If f is a continuous function defined on a closed*[2] *and bounded*[3] *set*[4] *$\mathcal{A} \subset \mathbb{R}^N$, then f attains an absolute maximum and absolute minimum value on $\mathcal{A}$.*

Figure 6.14 illustrates the theorem for $\mathcal{A} \subset \mathbb{R}$.



(a) all conditions met



(b) not continuous

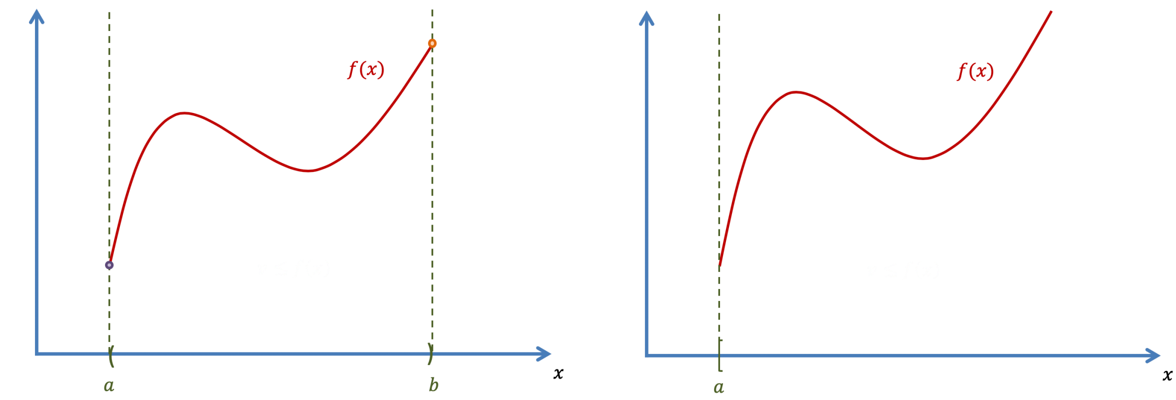
Figure 6.14: Extreme Value Theorem

[1] The qualification that at least one of the sets should have a non-empty interior rules out some awkward cases where the sets are of smaller dimension than the whole space. It is stated for completeness. In economic applications, the contour set of the objective function is full-dimensional, so no difficulty arises on this account.

[2] A set $\mathcal{A} \subset \mathbb{R}^N$ is *closed* if for every sequence $x^n \to x$ with $x^n \in \mathcal{A}$ for all n, we have $x \in \mathcal{A}$. That is, $\mathcal{A}$ is *closed* if it contains all its limit points.

[3] A set $\mathcal{A} \subset \mathbb{R}^N$ is *bounded* if there is $r \in \mathbb{R}$ such that $\|x\| < r$ for every $x \in \mathcal{A}$.

[4] A set $\mathcal{A} \subset \mathbb{R}^N$ that is both closed and bounded is also called *compact.*



(c) not closed (d) not bounded

Figure 6.14: Extreme Value Theorem (cont.)

In Figure 6.14a, all conditions in the theorem are met, and a maximum (minimum) is ensured. In the other three figures, one of the conditions are not met, and it is clear that a maximum (minimum) may not exist.

For our discussion, we impose the conditions that F and G are continuous, and that the constraint set is bounded. The continuity of G ensures the closedness of the set $\{x|G(x) \leq c\}$.[5]

Once we impose thses conditions, we could apply the Extreme Value Theorem and the existence of a maximum is ensured. Again, for our applications, these conditions are almost always satisfied and the existence of an optimum is not a problem. The conditions are stated here for the completeness of our discussion.

Separation. Let us return to our initial problem illustrated in Figure 6.1. Besides the conditions used to ensure the existence of a solution, we also impose the conditions that F is quasi-concave and G is quasi-convex, so that all conditions assumed in Figure 6.1 are met and the Seperation Theorem (Theorem 6.1) applies.

Let $px = b$ be the equation of the separating common tangent. The equation is unaffected if we multiply it through by -1, but will reverse the directions of the inequalities in (6.6). To ensure that the inequalities are consistent for the set $\mathcal{B}$ and $\mathcal{A}$ in Figure 6.1 and in Theorem 6.1, we choose $p_1, p_2 > 0$. The idea generalizes to more-variable cases.

[5] Appendix B provides a proof for this claim.

Since x^* lies on the separating tangent, so $px^* = b$. Therefore, (6.6) tells us that x^* gives the largest value of px among all points in $\mathcal{A}$, that is, among all points satisfying $G(x) \leq c$. Similarly, x^* gives the smallest value of px among all points in $\mathcal{B}$, that is, among all points satisfying $F(x) \geq v^*$. The result is summarized in Theorem 6.2 below.

Theorem 6.2 (Optimization by Separation)**.** *Given a quasi-concave function F and a quasi-convex function G, the point x^* maximizes $F(x)$ subject to $G(x) \leq c$ if, and only if, there is a non-zero vector p such that*

(i) x^* maximizes px subject to $G(x) \leq c$, and

(ii) x^* miminizes px subject to $F(x) \geq v^*$.

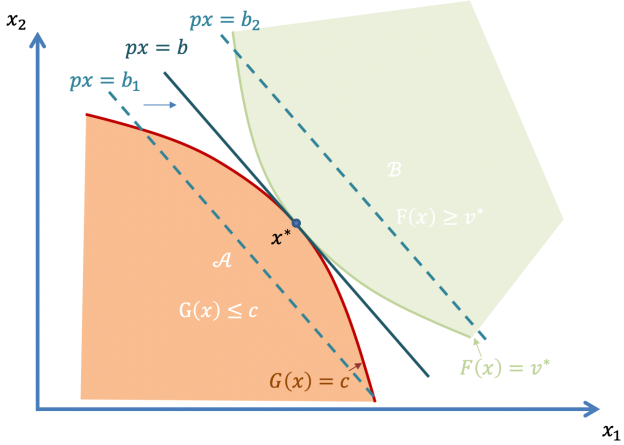
The generalization to several constraints is straightforward. The set $\mathcal{A}_i$ of points for which $G^i(x) \leq c_i$ is convex if G^i is quasi-convex. If this is so for all i, then the set $\mathcal{A}$ of points satisfying all the constraints, being the intersection of the convex sets $\mathcal{A}_i$, is also convex. Then Theorem 6.2 applies.

Note that Theorem 6.2 provides an "if and only if" result. That is, the conditions are both necessary and sufficient for optimality. But the problem with this theorem is that the conditions are not easy to verify in practical applications. In the next two chapters, we shall see sufficient conditions that are more useful in this regard.

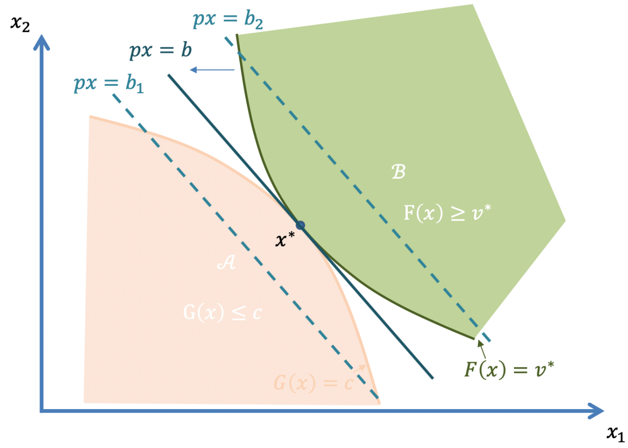
The real benefit from splitting the maximization problem into two separate problems comes from its economic interpretation. It raises the possibility of *decentralizing* optimal resource allocations using prices. Consider x as the production-cum-consumption vector, the constraints reflect limited resource availability, and the objective is the utility function. Now interpret p as the row vector of prices of outputs. The original problem of social optimization (Figure 6.1) can be decentralized. Apply Theorem 6.2,

(a) Part (i) says that the optimum x^* would be produced by a producer who seeks to maximize the value of output under the resource constraint. See Figure 6.15a.

(b) Part (ii) says that the optimum x^* would be consumed by a consumer who seeks to minimize the expenditure given the target utility level v^*. See Figure 6.15b.



(a) Producer Problem



(b) Consumer Problem

Figure 6.15: The Decentralization Problem

This separation of decision has two advantages:

i. Informational: the producer does not need to know the consumer's taste; and the consumer does not need to know the production technology.

ii. Incentives: the process relies on the self-interest of each side to ensure the effective implementation of the optimum.

Another remark is that only the relative prices matter for economic decisions. Nothing will change if we multiply the vector p and the related number b by the same positive number. This result is consistent with our discussions in the previous chapters.

The real life decentralization problem is more complicated. One problem is how the correct price vector is found, since people may not have the incentive to reveal their private information that is needed to calculate the right prices. Besides, issues of externality and distribution arise when there are many producers and consumers. These issues fall out of the scope of this course. Interested students may refer to microeonomic textbooks.

There exist situations where full decentralization is not possible. Recall Figure 6.13.

i. In case 6.13a, $\mathcal{B}$ is not convex and x^* does not minimize the expenditure in the consumer problem. The consumer prefers extremes to a diversified bundle of goods.

ii. In case 6.13b, $\mathcal{A}$ is not convex and x^* does not maximize the producer's value of output. The production technology has economies of scale or of specialization.

But in both cases, x^* does maximize $F(x)$ subject to $G(x) \leq c$. For x^* to be a maximizer of the original problem, what really matters is the *relative* curvature of F and G. We will discuss this idea and develop the conditions for maximization in Chapter 8.

6.D. Uniqueness

In Figure 6.1, the boundaries of the sets $\mathcal{B}$ and $\mathcal{A}$ are shown as smooth curves. But in general, a convex set can have straight-line segments along its boundary. Such possibilities have implications for separation and optimization. We consider the cases in Figure 6.16.

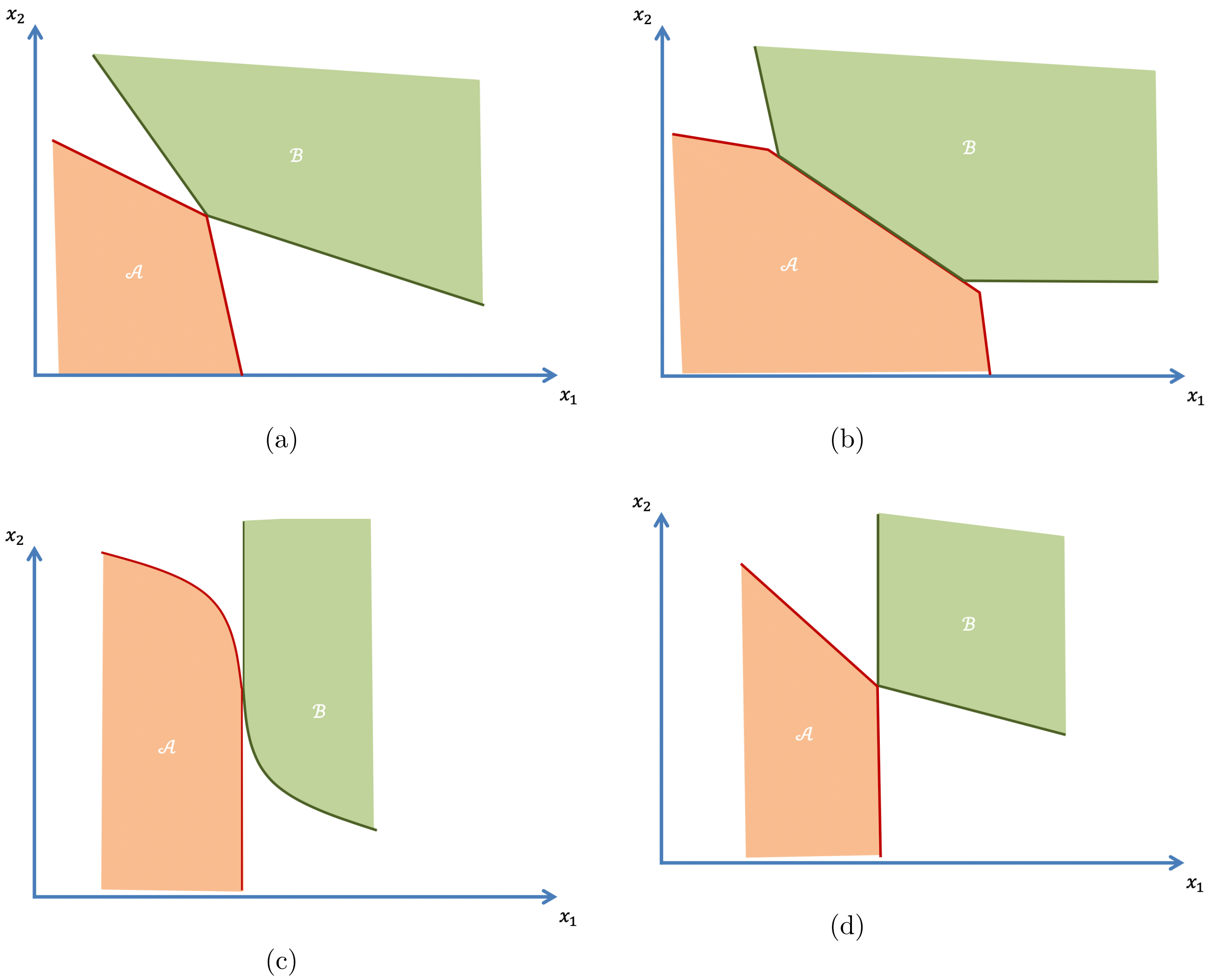


Figure 6.16: Optima at kinks and along flats

(a) In 6.16a, two corners happen to meet at x^*. Now we can find many lines through x^* that separate the two sets, that is, the decentralizing price vector p is not unique. It is not a serious problem for decentralization. In fact, the separation is a more general notion than that of a common tangent. And the decentralization depends on the separation property. Theorem 6.2 continues to hold.

(b) In 6.16b, the two sets have a flat portion in common. Now any points along this region serves as the optimum x^*. It causes problems about decentralization. Given p, all points on the flat portion of $\mathcal{A}$ will yield the same value of output to the producer; and all those on the flat portion of $\mathcal{B}$ will yield the same utility to the consumer. Thus, there is no reason to believe that the choices made separately by the producer and the consumer would coincide. In such a situation, we could only make a weaker claim: *if* the producer and the consumer happen to make coincident choices, neither will have strict incentive to depart from such choices.

(c) In 6.16c, the two boundaries have vertical parts in common. Then, we will have a vertical separating line, indicating $p_2 = 0$. In such cases, good 2 is a free good. Similarly, horizontal separating lines imply $p_1 = 0$, i.e., good 1 is a free good.

In 6.16d, there is a vertical separating line, whereas there are also non-vertical ones. Without stronger assumptions, it is not possible to guarantee strictly positive prices.

(d) The case of a positive slope of the common tangent, or negative price of either good, is usually avoided by assuming either that "free disposal" is possible so that the boundary of $\mathcal{A}$ cannot slope upward; or that both goods are desirable so that the boundary of $\mathcal{B}$ cannot slope upward. In our figures, these assumptions are implicit.

To summarize, problems of kinks are not serious. In fact, such cases generalize the concept of tangency and preserve the decentralization property. However, problems of flats are more serious because optimal choices can be non-unique and decentralization becomes problematic. We will discuss the additional assumptions needed to avoid this problem. In fact, a strengthening of the concepts of quasi-concavity (-convexity) will suffice.

The problem arises since the definition of convex sets allow straight-line segments along its boundary. To avoid the problem caused by the straight-line segments, we need to define the following stronger concepts. The idea is to modify the concept of convex sets to require all points of the line segment except the end-points to be interior points.

Definition 6.D.1 (Strongly Convex Set)**.** A set $\mathcal{S}$ of points in n-dimensional space is called *strongly convex* if, given any two points $x^a \in \mathcal{S}$ and $x^b \in \mathcal{S}$ and any real number $\alpha \in (0,1)$, the point $\alpha x^a + (1-\alpha)x^b$ is interior to $\mathcal{S}$.

Correspondingly, we modify the concepts of quasi-concavity and quasi-convexity.

Definition 6.D.2 (Strictly Quasi-concave Function)**.** A function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is *strictly quasi-concave* if the set $\{x|f(x) \geq c\}$ is strongly convex for all $c \in \mathbb{R}$, or equivalently, if

$$f(\alpha x^a + (1-\alpha)x^b) > \min\{f(x^a), f(x^b)\},$$

for all x^a, x^b and for all $\alpha \in (0,1)$.

Definition 6.D.3 (Strictly Quasi-convex Function)**.** A function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is *strictly quasi-convex* if the set $\{x|f(x) \leq c\}$ is strongly convex for all $c \in \mathbb{R}$, or equivalently, if

$$f(\alpha x^a + (1-\alpha)x^b) < \max\{f(x^a), f(x^b)\},$$

for all x^a, x^b and for all $\alpha \in (0,1)$.

Next, consider again the problem of maximizing $F(x)$ subject to $G(x) \leq c$, but now consider F being **strictly** quasi-concave, and G still being quasi-convex. Suppose x^* satisfies the conditions of Theorem 6.2. Then, x^* must be a unique solution.
To see this, we show by contradiction. Suppose that $\hat{x}$ is another solution. Then, x^* and $\hat{x}$ should be optimal for the consumer's problem. Thus, $px^* = p\hat{x} = b$ and $F(x^*) = F(\hat{x}) = v^*$. Now consider the point $\tilde{x} = \alpha x^* + (1-\alpha)\hat{x}$, for some $\alpha \in (0,1)$.

(i) $p\tilde{x} = p(\alpha x^* + (1-\alpha)\hat{x}) = \alpha px^* + (1-\alpha)p\hat{x} = \alpha b + (1-\alpha b) = b$.

(ii) Since F is strictly quasi-concave, $F(\tilde{x}) > \min\{F(x^*), F(\hat{x})\} = v^*$.

By continuity of F, there exists $\beta < 1$, such that $F(\beta\tilde{x}) > v^*$ (i.e., $\beta\tilde{x}$ is interior to $\mathcal{B}$). Besides, $p(\beta\tilde{x}) < p\tilde{x} = b$. Thus, the bundle $\beta\tilde{x}$ is interior to $\mathcal{B}$ with $p(\beta\tilde{x}) < b$, contradicting with the separation property. Therefore, the initial supposition must be wrong and strict quasi-concavity of F implies the uniqueness of the maximizer.

Remark. **Strict** quasi-convexity of G together with quasi-concavity of F also imply the uniqueness of the maximizer. If there are more than one constraint, we require every component constraint function G^i to be **strictly** quasi-convex.

6.E. Examples

Example 6.1: Illustration of Separation.

Consider the following problem:

$$\max_{x\geq 0,\, y\geq 0} F(x,y) = xy$$
$$\text{s.t. } G(x,y) = x^2 + y^2 \leq 25.$$

Figure 6.17 illustrates the separation.

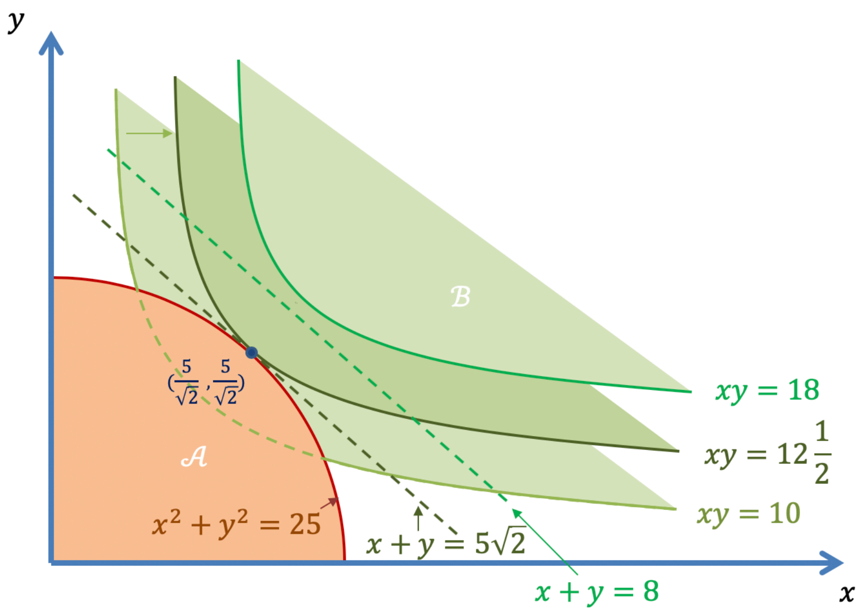


Figure 6.17: Illustration of Separation

The feasible set $\mathcal{A}$ consists of the quarter-circle (red) and all points below it (orange); boundaries of the upper contour sets of F for various values v are shown as a family of rectangular hyperbolas (green). The optimal occurs at $(x^*, y^*) = (5/\sqrt{2}, 5/\sqrt{2})$ and the maximized value of $F(x,y)$ is $v^* = 12\frac{1}{2}$.

i. The upper contour set $\mathcal{B}$ corresponding to v^* touches the feasible set at the optimum; they are separated by a common tangent $x + y = 5\sqrt{2}$.

ii. The upper contour set of F for the larger value 18 has no points in common with $\mathcal{A}$. We can draw a separating line $x + y = 8$ through the clear gap between them.

iii. For a smaller value than v^*, say 10, the upper contour set of F and the feasible set have interior points in common, and the two cannot be separated.

Example 6.2: Indirect Utility and Expenditure Functions.

Part I: Expenditure Function. The expenditure function is

$$E(p, u) = \min_x \{px | U(x) \geq u\}.$$

Show that $E(p, u)$ is concave in p for each fixed u.

Solution. To show concavity of $E(p, u)$ in p, we need to show that for any price vectors p^a and p^b and any number $\alpha \in [0, 1]$, we have

$$E(\alpha p^a + (1-\alpha)p^b, u) \geq \alpha E(p^a, u) + (1-\alpha)E(p^b, u) \tag{6.7}$$

Let x^c achieve the expenditure minimization for the price vector $\alpha p^a + (1-\alpha)p^b$, i.e.,

$$E(\alpha p^a + (1-\alpha)p^b, u) = \left(\alpha p^a + (1-\alpha)p^b\right) x^c.$$

Since x^c is feasible for the price vector $\left(\alpha p^a + (1-\alpha)p^b\right)$, x^c must satisfy the constraint, i.e., $U(x^c) \geq u$. The constraint does not involve the price vectors, so x^c is also feasible when the price vector is p^a or p^b. Then by the definition of $E(p, u)$,

$$E(p^a, u) \leq p^a x^c \quad \text{and} \quad E(p^b, u) \leq p^b x^c.$$

Multiply the first inequality by $\alpha \in [0, 1]$ and the second by $(1-\alpha)$, and adding the two, we have

$$\begin{aligned} \alpha E(p^a, u) + (1-\alpha)E(p^b, u) &\leq \alpha p^a x^c + (1-\alpha)p^b x^c \\ &= \left(\alpha p^a + (1-\alpha)p^b\right) x^c = E(\alpha p^a + (1-\alpha)p^b, u). \end{aligned}$$

This proves (6.7).

The economic intuition is illustrated in Figure 6.18 below. Consider the change of p_1 alone. As p_1 changes, one *could* leave the quantity vector unchanged. Then the expenditure would change linearly with the price. See the blue line. To the extent that there is substitution along the indifference curves, the quantity choice can be adapted to the changing prices. This will change the expenditure lower than linearly, that is, the minimized expenditure will be a concave function of prices. See the red curve.

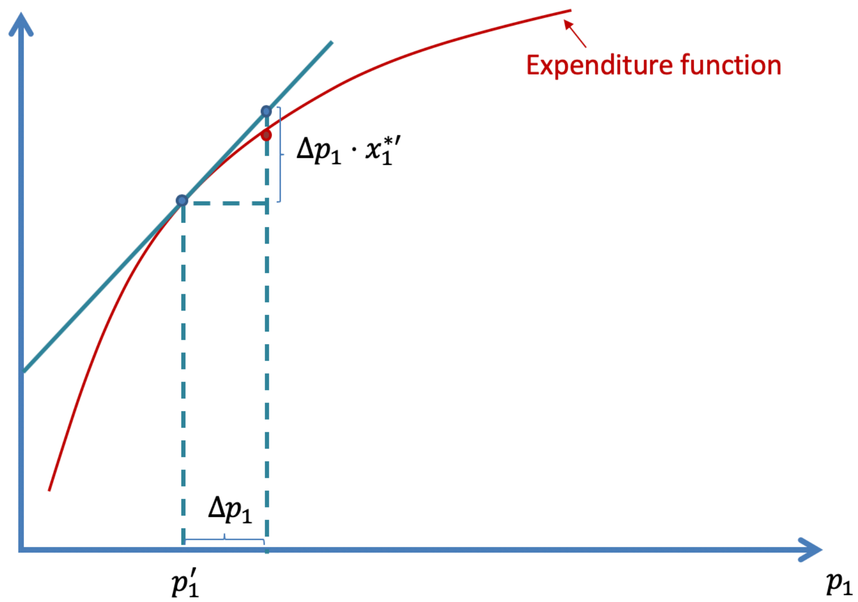


Figure 6.18: Expenditure Function

Part II: Indirect Utility Function. The indirect utility function is

$$V(p, I) = \max_x \{U(x)|px \leq I\}.$$

Show that $V(p, I)$ is quasi-convex in (p, I).

Solution. To show quasi-convexity of $V(p, I)$ in (p, I), we need to show that for any price-income vectors (p^a, I^a) and (p^b, I^b) and any number $\alpha \in [0, 1]$, we have

$$V\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right) \leq \max\left\{V\left(p^a, I^a\right), V\left(p^b, I^b\right)\right\} \tag{6.8}$$

Let x^c be the utility-maximizing bundle for $\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right)$. That is,

$$V\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right) = U(x^c)$$

Since x^c is feasible for the price-income vector $\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right)$, x^c must satisfy the constraint, that is

$$\left(\alpha p^a + (1-\alpha)p^b\right) x^c \leq \alpha I^a + (1-\alpha)I^b. \tag{6.9}$$

The proof goes in two steps. First, we show that x^c is feasible for at least one of the price-income vectors (p^a, I^a) and (p^b, I^b).

Suppose it is not, then we must have

$$p^a x^c > I^a \quad \text{and} \quad p^b x^c > I^b.$$

Multiply the first inequality by $\alpha \in [0,1]$ and the second by $(1-\alpha)$, and adding the two, we have

$$\alpha p^a x^c + (1-\alpha)p^b x^c > \alpha I^a + (1-\alpha)I^b \iff \left(\alpha p^a + (1-\alpha)p^b\right) x^c > \alpha I^a + (1-\alpha)I^b,$$

contradicting with (6.9). Therefore, x^c must be feasible for at least one of the price-income vectors (p^a, I^a) and (p^b, I^b).

Then, in whichever situation that x^c is feasible, by the definition of $V(p, I)$, $U(x^c)$ cannot exceed the maximum utility achievable in that situation. Therefore, at least one of

$$U(x^c) \leq V(p^a, I^a) \quad \text{and} \quad U(x^c) \leq V(p^b, I^b)$$

must hold. Therefore,

$$U(x^c) \leq \max\{V(p^a, I^a), V(p^b, I^b)\},$$

and thus (6.8) holds.

The intuition is illustrated in Figure 6.19 below. Here we consider only two goods. The area on or below the blue line denotes the feasible bundles under the price-income vector (p^a, I^a). The area on or below the yellow line denotes the feasible bundles under the price-income vector (p^b, I^b). The area on or below the green line denotes the feasible bundles under the price-income vector $\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right)$. It is clear that the area on or below the green line is either covered by the area on or below the blue line or covered by the area on or below the yellow line. Therefore, the optimal bundle under the price-income vector $\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right)$ (on or below the green line) is inside the union of the area on or below the blue line and the yellow line. Thus, such a bundle must be attainable under either the price-income vector (p^a, I^a) or the price-income vector (p^b, I^b). On the other hand, since the union of the areas are larger than the area on or below the green line, the better of the optimal bundle under the price-income vector (p^a, I^a) and that under the price-income vector (p^b, I^b) may possibly attain higher utility compared to the optimal bundle under the the price-income vector $\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right)$. That is,

$$V\left(\alpha p^a + (1-\alpha)p^b, \alpha I^a + (1-\alpha)I^b\right) \leq \max\left\{V\left(p^a, I^a\right), V\left(p^b, I^b\right)\right\}.$$

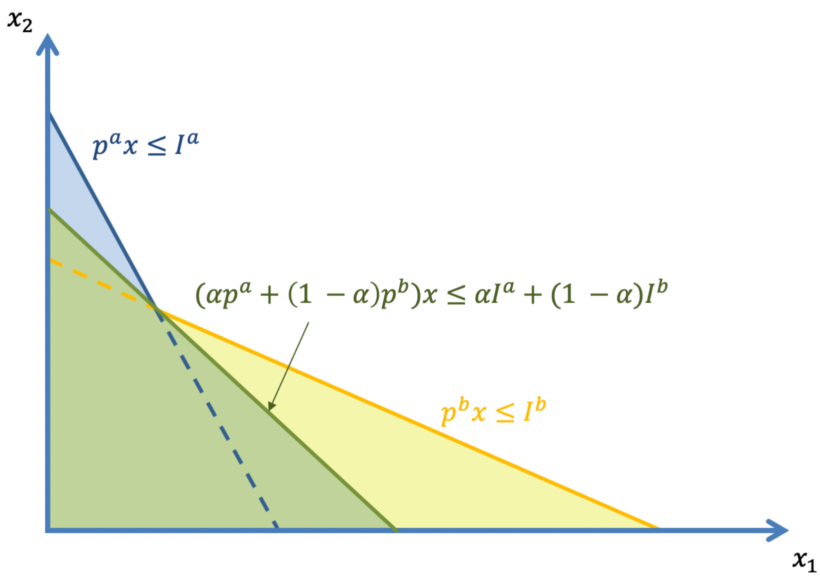


Figure 6.19: Budget Sets

Appendix A

Theorem 1 (Separating Hyperplane Theorem (Part I)). *Suppose that $\mathcal{B} \subset \mathbb{R}^N$ is convex and closed, and that $y \notin \mathcal{B}$. Then there is a $p \in \mathbb{R}^N$ with $p \neq 0$, and a value $c \in \mathbb{R}$ such that $p \cdot y > c$ and $p \cdot x < c$ for every $x \in \mathcal{B}$.*

Proof. For any $z \in \mathbb{R}^N$ and $z \neq y$, define $p = y - z$. First, $p \cdot y > p \cdot z$ because $p \cdot (y - z) = \|y - z\|^2 > 0$. Let $c = p \cdot \left(\frac{y+z}{2}\right)$ so that $p \cdot y > c > p \cdot z$.
Suppose $z = \arg\min_{x \in \mathcal{B}} \|y - x\|^2$ (see Figure 6.20). (Existence of z? Yes, but needs proof.[6]) Consider an arbitrary $x \in \mathcal{B}$.

$$\begin{aligned}
\|z-y\|^2 &\leq \|(1-\lambda)\, z + \lambda x - y\|^2 = \|(1-\lambda)\,(z-y) + \lambda\,(x-y)\|^2 \\
&= (1-\lambda)^2 \|z-y\|^2 + \lambda^2 \|x-y\|^2 + 2\,(1-\lambda)\,\lambda\,(z-y)\cdot(x-y) \\
\implies\ & 0 \leq \lambda\,(\lambda-2)\,\|z-y\|^2 + \lambda^2 \|x-y\|^2 + 2\,(1-\lambda)\,\lambda\,(z-y)\cdot(x-y) \\
\implies\ & 0 \leq (\lambda-2)\,\|z-y\|^2 + \lambda\,\|x-y\|^2 + 2\,(1-\lambda)\,(z-y)\cdot(x-y)
\end{aligned}$$

Taking limit, letting λ go to zero,

$$\begin{aligned}
&0 \leq -2\,(z-y)\cdot(z-y) + 2\,(z-y)\cdot(x-y) \\
\implies &0 \leq 2\,(z-y)\cdot(x-z) = -2p\cdot(x-z) \implies p\cdot z \geq p \cdot x.
\end{aligned}$$

Therefore, $p \cdot y > c > p \cdot z \geq p \cdot x$ for all $x \in \mathcal{B}$. □

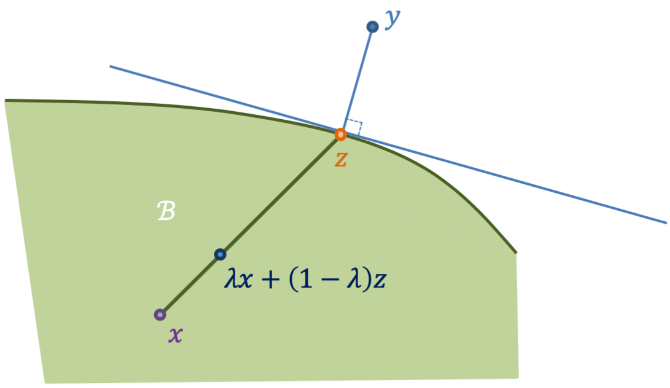


Figure 6.20: Separating Hyperplane Theorem (Part I)

[6] To prove the existence of z, we invoke the Extreme Value Theorem. However, we could not apply the theorem directly since the set $\mathcal{B}$ may not be bounded. Define a new set $\hat{\mathcal{B}} = \{x \in \mathcal{B} : \|q - y\| \geq \|x - y\|\}$ for some $q \in \mathcal{B}$ Then, $\hat{\mathcal{B}}$ is closed and bounded. $\hat{\mathcal{B}}$ is bounded since $\|x\| \underbrace{\leq}_{\text{Triangle Inequality}} \|y\| + \|y - x\| \underbrace{\leq}_{\text{Definition of } \hat{\mathcal{B}}} \|y\| + \|q - y\|$.
We could apply the *Extreme Value Thereom* on $\hat{\mathcal{B}}$ and get a minimizer z. For those $x \in \mathcal{B} \backslash \hat{\mathcal{B}}$, $\|z - y\|^2 \leq \|q - y\|^2 < \|x - y\|^2$. Therefore, z minimizes $\|x - y\|^2$ for all $x \in \mathcal{B}$.

Theorem 2 (Supporting Hyperplane Theorem). *Suppose that $\mathcal{B} \subset \mathbb{R}^N$ is convex and that x is not an element of the interior of set $\mathcal{B}$ ($x \notin \operatorname{Int} \mathcal{B}$). Then there is $p \in \mathbb{R}^N$ with $p \neq 0$ such that $p \cdot x \geq p \cdot y$ for every $y \in \mathcal{B}$.*

Proof. Consider $x \notin \operatorname{Int} \mathcal{B}$. Then we can find a sequence $x^m \to x$ such that for all m, x^m is not an element of the *closure*[7] of the set $\mathcal{B}$ ($x^m \notin \operatorname{Cl} \mathcal{B}$). By Theorem 1 Separating Hyperplane Theorem (Part I), for each m there is a $p^m \neq 0$ and a $c^m \in \mathbb{R}$ such that

$$p^m \cdot x^m > c^m \geq p^m \cdot y \tag{6.10}$$

for every $y \in \mathcal{B}$. Without loss of generality, suppose that $\|p^m\| = 1$ for every m. Thus, extracting a subsequence if necessary[8], we can assume that there is $p \neq 0$ and $c \in \mathbb{R}$ such that $p^m \to p$ and $c^m \to c$. Taking limits of 6.10, we have

$$p \cdot x \geq c \geq p \cdot y$$

for every $y \in \mathcal{B}$. □

Theorem 3 (Separating Hyperplane Theorem (Part II)). *Suppose that the convex sets $\mathcal{A}, \mathcal{B} \subset \mathbb{R}^N$ are disjoint (i.e., $\mathcal{A} \cap \mathcal{B} = \varnothing$). Then there is $p \in \mathbb{R}^N$ with $p \neq 0$, and a value $c \in \mathbb{R}$, such that $p \cdot x \geq c$ for every $x \in \mathcal{A}$ and $p \cdot y \leq c$ for every $y \in \mathcal{B}$. That is, there is a hyperplane that separates $\mathcal{A}$ and $\mathcal{B}$, leaving $\mathcal{A}$ and $\mathcal{B}$ on different sides of it.*

Proof. Consider arbitrary $x \in \mathcal{A}$ and $y \in \mathcal{B}$ and let $z = x - y$. Let

$$\mathcal{D} = \left\{ z \in \mathbb{R}^N : z = x - y \text{ for some } x \in \mathcal{A} \text{ and some } y \in \mathcal{B} \right\}.$$

Now we show that $\mathcal{D}$ is convex. Suppose $z_1, z_2 \in \mathcal{D}$. Then

$$\alpha z_1 + (1 - \alpha) z_2 = [\alpha x_1 + (1 - \alpha) x_2] - [\alpha y_1 + (1 - \alpha) y_2].$$

Since $\mathcal{A}$ and $\mathcal{B}$ are convex, $\alpha x_1 + (1 - \alpha) x_2 \in \mathcal{A}$ and $\alpha y_1 + (1 - \alpha) y_2 \in \mathcal{B}$.

So $\alpha z_1 + (1 - \alpha) z_2 \in \mathcal{D}$. Therefore, $\mathcal{D}$ is convex. Since $\mathcal{A}$ and $\mathcal{B}$ are disjoint, $0 \notin \mathcal{D}$. Since $0 \notin \mathcal{D}$, we have $0 \notin \operatorname{Int} \mathcal{D}$. Then, we could apply Thereom 2 Supporting Hyperplane

[7] A *closure* of a set $\mathcal{A}$ is the union of the set $\mathcal{A}$ and its limit points.

[8] The existence of convergent subsequence is a result of the Bolzano–Weierstrass Theorem: each bounded sequence in $\mathbb{R}^N$ has a convergent subsequence.

Theorem: there is $p' \in \mathbb{R}^N$ with $p' \neq 0$ such that $p' \cdot 0 \geq p' \cdot z$ for all $z \in \mathcal{D}$. Let $p = -p'$, we have $0 \leq p \cdot (x - y)$ or $p \cdot y \leq p \cdot x$ for all $x \in \mathcal{A}$ and $y \in \mathcal{B}$. To complete the proof, let

$$c = \frac{\inf_{x \in \mathcal{A}} p \cdot x + \sup_{y \in \mathcal{B}} p \cdot y}{2}. \quad \square$$

Theorem 6.1 (Separation Theorem). *If $\mathcal{B}$ and $\mathcal{A}$ are two convex sets, that have no interior points in common, and at least one of the sets has a non-empty interior,*[9] *then we can find a non-zero vector p and a number b such that the hyperplane $px = b$ separates the two sets, or*

$$px = \begin{cases} \leq b & \text{for all } x \in \mathcal{A} \\ \geq b & \text{for all } x \in \mathcal{B}. \end{cases} \tag{6.6}$$

Proof. Since the two convex sets $\mathcal{B}$ and $\mathcal{A}$ have no interior points in common, we could apply Theorem 3 Separating Hyperplane Theorem (Part II) to $\operatorname{Int}\mathcal{B}$ and $\operatorname{Int}\mathcal{A}$. The boundaries of $\mathcal{B}$ and $\mathcal{A}$ are the limit points of the interiors, and thus must lie in the same half-space.[10] Therefore, the hyperplane separates the convex sets $\mathcal{B}$ and $\mathcal{A}$. $\square$

Appendix B

Claim. The continuity of G ensures the closedness of the set $\{x | G(x) \leq c\}$.

Proof. For closedness, we need to show that for any sequence $\{x^n\}_{n=1}^{\infty}$ where $G(x^n) \leq c$ for all n and $x = \lim_{n \to \infty} x^n$, we must have $G(x) \leq c$.
We prove by contradiction. Suppose that there exists a sequence $\{x^n\}_{n=1}^{\infty}$ where $G(x) \leq c$ for all n and $x = \lim_{n \to \infty} x^n$, but $G(x) > c$. Then by continuity of G, there exists $\varepsilon > 0$ such that for all y satisfying $\|y - x\| < \varepsilon$, we have $G(y) > c$. Therefore, $\exists N > 0$, s.t. $\forall n \geq N$, $G(x^n) > c$. This contradicts with $G(x^n) \leq c$ for all n. $\square$

[9] The qualification that at least one of the sets should have a non-empty interior rules out some awkward cases where the sets are of smaller dimension than the whole space. It is stated for completeness. In economic applications, the contour set of the objective function is full-dimensional, so no difficulty arises on this account.

[10] The idea is similar to the proof in Appendix B. Suppose that the boundary lies in the other half-space, then there must be some interior points in the other half-space, which leads to a contradiction.