Lecture 3: Convexity

Advanced Microeconomics I, ITAM

Xinyang Wang*

Convexity is prevalent in economic theory and optimizations. To give a few examples, an agent's preference usually exhibits some concavity; equilibrium price can be understood as the normal direction of some separate hyperplane of two disjoint convex sets; the interior minimizer of a convex function is conveniently characterized by the first order condition. In this section, we clarify the various concepts related to convexity.

It is worth noting that convexity is a concept directly related to the linear structure of the underlining space. Although in this notes, we will restrict our attention to the convexity in $\mathbb{R}^n$, we do not make use of any topological argument. In particular, a convex set may not have a "smooth" boundary, and a convex function may not be differentiable at some point.

1 Convex Sets

1.1 Definition and Basic Properties

A non-empty set $S \subset \mathbb{R}^n$ is *convex* if for all $x, y \in S$, $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in S$.

Graphically, a set is convex if, for any two points in the set, the line segment connecting them is in the set.

Example.

- *The unit ball* $B_1(0) = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ *is convex in* $\mathbb{R}^n$
- $B_1(0) - 0 = \{x \in \mathbb{R}^n : \|x\| \leq 1, x \neq 0\}$ *is not convex in* $\mathbb{R}^n$
- *The union of two disjoint balls is not convex*

*Please email me at xinyang.wang@itam.mx for typos or mistakes. Version: January 26, 2021.

Remark. *We say non-empty set $S \subset \mathbb{R}^n$ is mid-point convex if for any two points in the sets, their midpoint is in the set. i.e. for all $x, y \in S$, $\frac{1}{2}x + \frac{1}{2}y \in S$. While a convex set is clearly mid-point convex, the converse is not true. For instance, $\mathbb{Q} \cap [0,1]$ is mid-point convex in $\mathbb{R}$, but is not convex. Therefore, in the definition of convexity, we will need the whole line segment in the set. That is, the part about $\lambda \in [0,1]$ is indispensable.*

Proposition.

- *The intersection of convex sets is convex.*
- $S + T = \{s + t : s \in S, t \in T\}$ *is convex if* S, T *are convex*
- $\forall \lambda \geq 0$, $\lambda S = \{\lambda s : s \in S\}$ *is convex*

Proof. Exercise. ∎

1.2 Convex Hull

Next, we describe how to make a non-convex set convex by adding the missing part.

Given $x_1, ..., x_n \in \mathbb{R}^n$, we say x is a *convex combination* of $x_1, ..., x_n$ if

$$x = \sum_{i=1}^{n} a_i x_i$$

for some $a_i \in \mathbb{R}, \sum_{i=1}^{n} a_i = 1$.

Proposition. *For a convex set S and vectors $x_1, ..., x_n$ in S, any convex combination of $x_1, ..., x_n$ is in S.*

Proof. Exercise. (Hint: Prove by induction.) ∎

Next, we define convex hull, which is the set containing all convex combinations of points in the set. Formally, given a nonempty set $S \subset \mathbb{R}^n$, the convex hull of S, denoted by $co(S)$ is defined by the set of the convex combinations of finitely many points in S.

It is instant from the definition that $S \subset co(S)$, as a point is always a convex combination of itself. Moreover, an equivalent way to describe the convex hull of a set is the smallest convex set containing the set:

Theorem. *For any nonempty $S \subset \mathbb{R}^n$,*

$$co(S) = \bigcap_{S \subset G, G\, convex} G$$

Proof. Exercise. ∎

Corollary. *For any nonempty $S \subset \mathbb{R}^n$, S is convex if and only if $co(S) = S$.*

1.3 Separating Hyperplane Theorem

On the level of sets and linear spaces, one important property of convex sets are represented by the following graph:

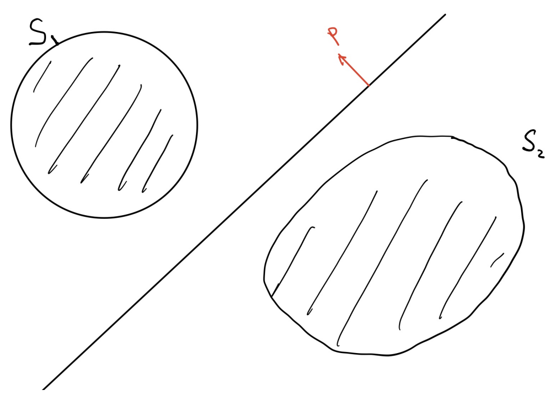

Figure 1: An Illustration of The Separating Hyperplane

That is, any two disjoint convex sets can be separated by a hyperplane. Now the question is how to write this idea formally. First, we we define hyperplane,

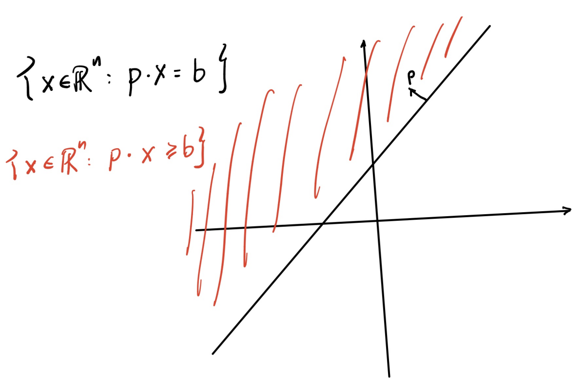

Figure 2: An Illustration of an Hyperplane

An *hyperplane* is defined by the set

$$\{x \in \mathbb{R}^n : p \cdot x = b\}$$

for some $p \in \mathbb{R}^n$ and $b \in \mathbb{R}$. Here, p is called the *normal vector* of the hyperplane, and b determines the translation of the hyperplane from the origin. The region above the the hyperplane can be represented by the set

$$\{x \in \mathbb{R}^n : p \cdot x \geq b\}$$

Here, one trick to determine the sign of number b is by checking whether the origin is in the shaded area. In the above picture, 0 is not in the red shaded area. Therefore, $p \cdot 0 \geq b$ should not holds. That is, b should be a positive number.

Having had a definition of hyperplane, it is clear that to write a set S is in the red shaded area in Figure 2, it is equivalent to say

$$S \subset \{x \in \mathbb{R}^n : p \cdot x \geq b\}$$

or, equivalently,

$$p \cdot x \geq b, \forall x \in S$$

Now, we are ready to write down our theorem:

Theorem (Separating Hyperplane)**.** *Let S_1, S_2 be two convex and disjoint in $\mathbb{R}^n$, then, there exists a normal vector $p \in \mathbb{R}^n$, $p \neq 0$, and a translation $b \in \mathbb{R}$ such that*

$$p \cdot x \geq b, \forall x \in S_1$$

$$p \cdot y \leq b, \forall y \in S_2$$

Remark. *In the statement of the theorem,*

1. *the convexity of two sets are indispensable. (Exercise)*

2. *the disjointness of two sets can be replaced by the disjointness of the relative interior of two sets. That is, intuitively, the two sets can touch on the boundary.*

3. *Under the current assumption, we can not replace any inequality by a strict inequality. (See Figure 3 below for a counter example)*

4. *When we impose an additional assumption that one of the convex sets is bounded, we can replace one inequality by a strict inequality.*

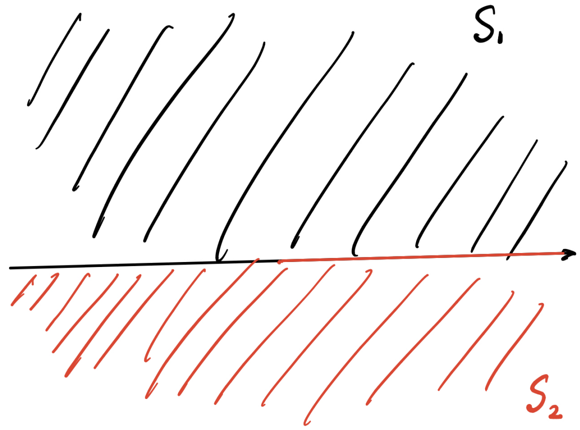

Figure 3: No strict inequality if no boundedness

The separating hyperplane theorem have numerous applications in economic theory. To give a few examples: as we will see later in this course, the normal vector will give us the equilibrium price in a competitive environment. Moreover, the separating hyperplane theorem can be used to prove Farkas' Lemma, which is closely related to the no-arbitrage condition in financial economics. (Probably in the next sequence of this course.)

2 Convex Functions

Given a convex set X in $\mathbb{R}^n$, we say a function $f : X \to \mathbb{R}$ is *convex* if

$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y), \forall x, y \in \mathbb{R}^n$$

In contrast, a function f is *concave* if $-f$ is convex.

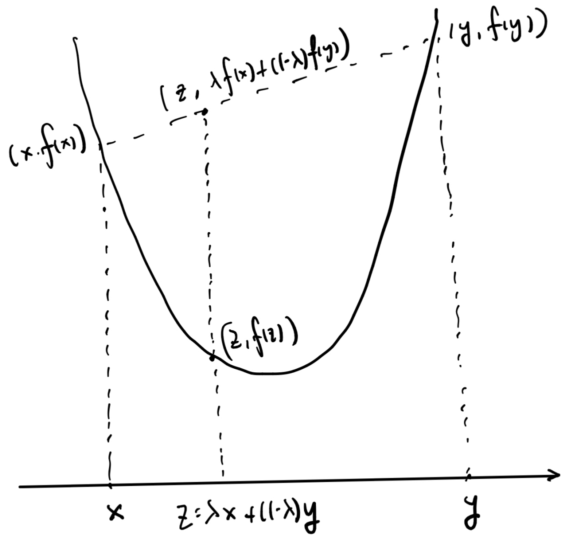

Figure 4: Illustration of a convex function

For those who have learned calculus, a natural question might be whether we could use the fact $f'' \geq 0$ to define a convex function. (note we have already know the meaning that a matrix is nonnegative) The answer is no, as we will use any topological structure of the domain X. In other words, the function may not be differentiable. (c.f. the utility function may not be differentiable in this course.)

2.1 Epigraph

Now, we wish to understand whether there is a relation between the convexity of a set and the convexity of a function. The answer is assertive. Indeed, there is a one to one relationship between the convexity of sets and convexity of functions in $\mathbb{R}^n$. In this part, we explore on this relation.

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, we define the *graph* of the function to be

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R} : y = f(x)\}$$

and we define the *epigraph* of the function to be

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R} : y \geq f(x)\}$$

That is, intuitively, the part above this function.

The following theorem represents the one to one relationship between the convexity of sets and the convexity of functions.

Theorem. *A function is convex if and only if its epigraph is a convex set.*

Proof. Exercise. ■

2.2 Another characterization of convex functions

It is valid to ask that why do we need to understand a convex function as a convex set. In this subsection, we give one application of this relation.

Theorem. *A function f is convex if and only if it is the supremum of a collection of affine functions.*[1] *i.e.*

$$f(x) = \sup_{\alpha \in \mathcal{A}} \{p_\alpha \cdot x + b_\alpha\}$$

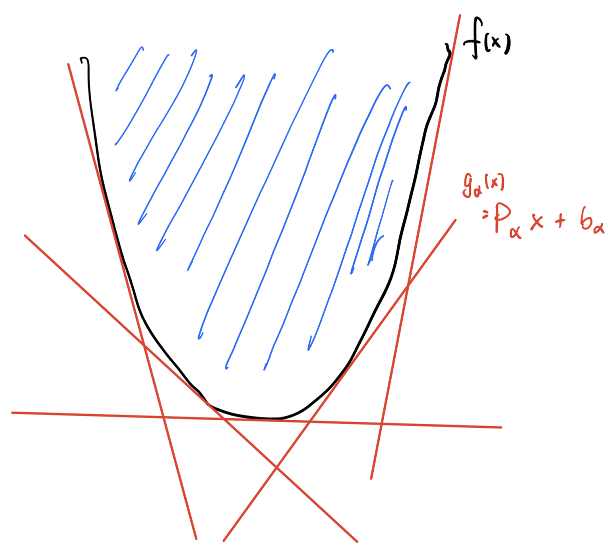

Figure 5: Supremum of a class of affine functions

Proof. (A Sketch of the proof) It is an exercise to prove that the supremum of affine functions is convex.

[1] Recall that an affine function on $\mathbb{R}^n$ is in the form $g(x) = p \cdot x + b$.

Conversely, given a convex function f, we know its epigraph (shaded in blue) is a convex set. Then for any point on the boundary of the set, the point has no intersection with the relative interior of the epigraph. Therefore, we apply the separating hyperplane theorem, and we argue the hyperplane must pass through the point. For each point on the boundary, we have an affine function $g_\alpha(x) = p_\alpha \cdot x + b_\alpha$. And the equation

$$f(x) = \sup_{\alpha \in \mathcal{A}} \{p_\alpha \cdot x + b_\alpha\}$$

is followed by the definition of the separating hyperplanes. ■