Chapter 3. Convex Sets and Concave Functions

Convexity is one of the most important mathematical properties in economics. For example, without convexity of preferences, demand and supply functions are not continuous, and so competitive markets generally do not have equilibrium points. The economic interpretation of convex preference sets in consumer theory is diminishing marginal rates of substitution; the interpretation of convex production sets is constant or decreasing returns to scale. Considerably less is known about general equilibrium models that allow non-convex production sets (e.g., economies of scale) or non-convex preferences (e.g., the consumer prefers a pint of beer or a shot of vodka alone to any mixture of the two). We refer to Nikaido (1968), Rockafellar (1970) and Boyd and Vandenberghe (2004) for more general discussions on convexity; see Hiriart-Urruty and Lemaréchal (2001) for an introduction.

The emphasis of this chapter is to show uniqueness of the optimizer and sufficient conditions for optimization through convexity. Related materials can be found in Chapter 21 of Simon and Blume (1994) and Chapter 7-8 of Sundaram (1996). In this chapter, lowercase bold letters such as \( x = (x_1, \ldots, x_n)' \) represent column vectors.

1 Convex Sets

1.1 Basics

Given two points \( x, y \in \mathbb{R}^n \), a point \( z = tx + (1 - t)y \), where \( 0 \leq t \leq 1 \), is called a convex combination of \( x \) and \( y \). The set of all possible convex combinations of \( x \) and \( y \), denoted by \([x, y]\), is called the interval with endpoints \( x \) and \( y \) (or, the line segment connecting \( x \) and \( y \)), i.e.,

\[
[x, y] = \{tx + (1 - t)y \mid 0 \leq t \leq 1\}.
\]

This definition is an extension of the interval in \( \mathbb{R}^1 \).

**Definition 1** A set \( S \subseteq \mathbb{R}^n \) is convex iff for any points \( x \) and \( y \) in \( S \) the interval \([x, y] \subseteq S\).

In words: a set is convex if it contains the line segment connecting any two of its points; or, more loosely speaking, a set is convex if for any two points in the set it also contains all points between them.

*Email: pingyu@hku.hk*
Convex sets in $\mathbb{R}^2$ include interiors of triangles, squares, circles, ellipses, and hosts of other sets. Note also that, for example in $\mathbb{R}^3$, while the interior of a cube is a convex set, its boundary is not. (Of course, the same is true of the square in $\mathbb{R}^2$.) The quintessential convex set in Euclidean space $\mathbb{R}^n$ for any $n > 1$ is the $n$-dimensional open ball $B_r(a)$ of radius $r > 0$ about point $a \in \mathbb{R}^n$, where recall from Chapter 1 that

$$B_r(a) = \{ x \in \mathbb{R}^n \mid \| x - a \| < r \}.$$ 

Example 1 Prove that the budget constraint $B = \{ x \in X : p'x \leq y \}$ is convex.

Proof. For any two points $x_1, x_2 \in B$, we have $p'x_1 \leq y$ and $p'x_2 \leq y$. Then for any $t \in [0, 1]$, we must have $tx_1 + (1 - t)x_2 \leq y$. This is equivalent to say that $tx_1 + (1 - t)x_2 \in B$. So the budget constraint $B$ is convex.

Exercise 1 Is the empty set convex? Is a singleton convex? Is $\mathbb{R}^n$ convex? In each case prove that the set is convex or prove that it is not.

There are also several standard ways of forming convex sets from convex sets. We now examine a number of such ways.

Let $A, B \subseteq \mathbb{R}^n$ be sets. The Minkowski sum $A + B \subseteq \mathbb{R}^n$ is defined as

$$A + B = \{ x + y \mid x \in A, y \in B \}. $$

When $B = \{ b \}$ is a singleton, the set $A + \{ b \}$ is often written $A + b$ is called a translation of $A$.

Exercise 2 Prove that if $A$ and $B$ are convex then $A + B$ is convex.

Let $A \subseteq \mathbb{R}^n$ be a set and $\alpha \in \mathbb{R}$ be a number. The scaling $\alpha A \subseteq \mathbb{R}^n$ is defined as

$$\alpha A = \{ \alpha x \mid x \in A \}. $$

When $\alpha > 0$, the set $\alpha A$ is called a dilation of $A$.

Exercise 3 Prove that if $A$ is convex then for any $\alpha \in \mathbb{R}$ the set $\alpha A$ is convex.

Exercise 4 Prove that for any finite number of convex sets $S_1, \ldots, S_K$ the intersection $\cap_{i=1}^K S_i$ is convex. In fact the result does not depend on the number of sets being finite. Prove that the intersection $\cap_{i \in I} S_i$ of any number of convex sets is convex.

Exercise 5 Show by example that the union of a number of convex sets need not be convex.

It is also possible to define the convex combination of an arbitrary (but finite) number of points.
Definition 2 Let $x_1, ..., x_k$ be a finite set of points from $\mathbb{R}^n$. A point

$$x = \sum_{i=1}^{k} \alpha_i x_i,$$

where $\alpha_i \geq 0$ for $i = 1, ..., k$ and $\sum_{i=1}^{k} \alpha_i = 1$, is called a **convex combination** of $x_1, ..., x_k$.

Note that the definition of a convex combination of two points is a special case of this definition.

Can we generate ‘superconvex’ sets using Definition 2? That is, if we start with a convex set can we make it even more convex by adding all the convex combinations of points in the set? No, as the following Lemma shows.

Lemma 1 A set $S \subseteq \mathbb{R}^n$ is convex iff every convex combination of points of $S$ is in $S$.

**Proof.** If a set contains all convex combinations of its points it is obviously convex, because it also contains convex combinations of all pairs its points. Thus, we need to show that a convex set contains any convex combination of its points. The proof is by induction on the number of points of $S$ in a convex combination. By definition, convex set contains all convex combinations of any two of its points. Suppose that $S$ contains any convex combination of $n$ or fewer points and consider one of $n+1$ points, $x = \sum_{i=1}^{n+1} \alpha_i x_i$. Since not all $\alpha_i = 1$, we can relabel them so that $\alpha_{n+1} < 1$. Then

$$\begin{align*}
x &= (1 - \alpha_{n+1}) \sum_{i=1}^{n} \frac{\alpha_i}{1 - \alpha_{n+1}} x_i + \alpha_{n+1} x_{n+1} \\
&= (1 - \alpha_{n+1}) y + \alpha_{n+1} x_{n+1}.
\end{align*}$$

Note that $y \in S$ by induction hypothesis (as a convex combination of $n$ points of $S$) and, as a result, so is $x$, being a convex combination of two points in $S$. \(\blacksquare\)

But, using Definition 2, we can generate convex sets from non-convex sets. This operation is very useful, so the resulting set deserves a special name.

Definition 3 Given a set $S \subseteq \mathbb{R}^n$, the set of all convex combinations of points from $S$, denoted $\text{conv} S$, is called the **convex hull** of $S$.

Exercise 6 Prove that for any set $S \subseteq \mathbb{R}^n$ the convex hull of $S$ is a convex set. (This is not difficult. It just involves a careful attention to the details of the definition. Or, put another way, the only difficulty is seeing that there is something to prove.)

In light of the previous exercise, Lemma 1 can be written more succinctly as $S = \text{conv} S$ iff $S$ is convex.
1.2 Convex Hulls (*)

The next theorem deals with the following interesting property of convex hulls: the convex hull of a set $S$ is the intersection of all convex sets containing $S$. Thus, in a natural sense, the convex hull of a set $S$ is the ‘smallest’ convex set containing $S$. In fact, many authors define convex hulls in that way and then prove our Definition 3 as theorem.

**Theorem 1** Let $S \subseteq \mathbb{R}^n$ be a set. Then any convex set containing $S$ also contains $\text{conv } S$.

**Proof.** Let $A$ be a convex set such that $S \subseteq A$. By lemma 1, $A$ contains all convex combinations of its points and, in particular, all convex combinations of points of its subset $S$, which is $\text{conv } S$.

The next exercise is again quite obvious. It again, frustrates attempts to generate ‘superconvex’ sets, this time by showing that we do not make a set more convex by taking the convex hull of a convex hull.

**Exercise 7** Prove that $\text{conv conv } S = \text{conv } S$ for any $S$.

**Exercise 8** Prove that if $A \subseteq B$ then $\text{conv } A \subseteq \text{conv } B$.

The next asks you to show that when taking convex hulls and taking direct sums, it does not matter in which order you use these operations.

**Exercise 9** Prove that $\text{conv } (A + B) = (\text{conv } A) + (\text{conv } B)$.

On the other hand when taking unions or intersections and convex hulls the order may matter.

**Exercise 10** Prove that $\text{conv } (A \cap B) \subseteq (\text{conv } A) \cap (\text{conv } B)$. Give an example to show that the inclusion may be strict.

**Exercise 11** Prove that $\text{conv } (A \cup (\text{conv } B) \subseteq \text{conv } (A \cup B)$. Again, give an example to show that the inclusion may be strict.

1.3 Caratheodory’s Theorem (*)

Definition 3 implies that any point $x$ in the convex hull of $S$ is representable as a convex combination of (finitely) many points of $S$ but it places no restrictions on the number of points of $S$ required to make the combination. Caratheodory’s Theorem puts the upper bound on the number of points required - in $\mathbb{R}^n$ the number of points never has to be more than $n + 1$.

**Theorem 2 (Caratheodory, 1907)** Let $S \subseteq \mathbb{R}^n$ be a non-empty set; then every $x \in \text{conv } S$ can be represented as a convex combination of (at most) $n + 1$ points from $S$.

Note that the theorem does not ‘identify’ the points used in the representation, their choice would depend on $x$.

**Exercise 12** Show by example that the constant $n + 1$ in Caratheodory’s theorem cannot be improved. That is, exhibit a set $S \subseteq \mathbb{R}^n$ and a point $x \in \text{conv } S$ that cannot be represented as a convex combination of fewer than $n + 1$ points from $S$. 

4
1.4 Polytopes (*)

The simplest convex sets are those which are convex hulls of a finite set of points, i.e., sets of the form \( S = \text{conv}\{x^1, x^2, ..., x^m\} \). The convex hull of a finite set of points in \( \mathbb{R}^n \) is called a polytope.

**Exercise 13** Prove that the set

\[ \Delta = \{ x \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i = 1 \text{ and } x_i \geq 0 \text{ for any } i \} \]

is a polytope. This polytope is called the **standard \( n \)-dimensional simplex**.

**Exercise 14** Prove that the set

\[ C = \{ x \in \mathbb{R}^n \mid 0 \leq x_i \leq 1 \text{ for any } i \} \]

is a polytope. This polytope is called an **\( n \)-dimensional cube**.

**Exercise 15** Prove that the set

\[ O = \{ x \in \mathbb{R}^n \mid \sum_{i=1}^{n} |x_i| \leq 1 \} \]

is a polytope. This polytope is called a (**hyper**)octahedron.

1.5 The Topology of Convex Sets (*)

We have now looked at another structure on \( \mathbb{R}^n \) along with the topological and algebraic structures we looked at earlier. The following result tells us something about how these structures are related.

**Proposition 1** The closure of a convex set is a convex set. The interior of a convex set (possible empty) is convex.

Recall that you showed earlier that the empty set was convex.

2 Convex Functions

In order to be able to say whether or not the optimization problem has a unique solution it is useful to know something about the shape or **curvature** of the functions \( f \) and \((g, h)\). Also, the curvature of \( f \) and \((g, h)\) is related to whether the Kuhn-Tucker conditions are sufficient to identify the maximizers.

2.1 Basics

We say a function is **concave** if for any two points in the domain of the function the value of function at a weighted average of the two points is greater than the weighted average of the values of the function at the two points. We say the function is **convex** if the value of the function at the average is less than the average of the values. The following definition makes this a little more explicit.
Definition 4 A function \( f : S \rightarrow \mathbb{R} \) defined on a convex set \( S \) is concave if for any \( x, x' \in S \) with \( x \neq x' \) and for any \( t \) such that \( 0 < t < 1 \) we have \( f(tx + (1-t)x') \geq tf(x) + (1-t)f(x') \). The function is strictly concave if \( f(tx + (1-t)x') > tf(x) + (1-t)f(x') \).

A function \( f : S \rightarrow \mathbb{R} \) defined on a convex set \( S \) is convex if for any \( x, x' \in S \) with \( x \neq x' \) and for any \( t \) such that \( 0 < t < 1 \) we have \( f(tx + (1-t)x') \leq tf(x) + (1-t)f(x') \). The function is strictly convex if \( f(tx + (1-t)x') < tf(x) + (1-t)f(x') \).

Exercise 16 Why don’t we check \( t = 0 \) and \( 1 \) in the definition? Why the domain of \( f \) must be a convex set?

Exercise 17 We say that the function \( f(x_1, \ldots, x_n) \) is nondecreasing if \( x'_i \geq x_i \) for each \( i \) implies that \( f(x'_1, \ldots, x'_n) \geq f(x_1, \ldots, x_n) \), is increasing if \( x'_i > x_i \) for each \( i \) implies that \( f(x'_1, \ldots, x'_n) > f(x_1, \ldots, x_n) \) and is strictly increasing if \( x'_i \geq x_i \) for each \( i \) and \( x'_j > x_j \) for at least one \( j \) implies that \( f(x'_1, \ldots, x'_n) > f(x_1, \ldots, x_n) \). Show that if \( f \) is nondecreasing and strictly concave then it must be strictly increasing. [Hint: This is very easy.]

The following two figures show a typical concave and convex function intuitively. Note that although there are both concave and convex functions, there are only convex sets, no concave sets!

Remark 1 We can also define the concave/convex function through convex sets. For this purpose, we need to define the subgraph and epigraph of \( f \), denoted as \( \text{sub}(f) \) and \( \text{epi}(f) \):

\[
\text{sub}(f) = \{(x, y) \in S \times \mathbb{R} | f(x) \geq y\} \subset \mathbb{R}^{n+1},
\]

\[
\text{epi}(f) = \{(x, y) \in S \times \mathbb{R} | f(x) \leq y\} \subset \mathbb{R}^{n+1}.
\]
Figure 3 illustrates \( \text{sub}(f) \) and \( \text{epi}(f) \) intuitively. A function \( f \) is concave iff \( \text{sub}(f) \) is a convex set, and is convex iff \( \text{epi}(f) \) is convex.

In practice, we need some calculus criteria for concavity and convexity.

**Theorem 3** Let \( f \in C^2(U) \), where \( U \subseteq \mathbb{R}^n \) is open and convex. Then \( f \) is concave iff the Hessian

\[
D^2 f(x) = \begin{pmatrix}
\frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2}
\end{pmatrix}
\]

is negative semidefinite for all \( x \in U \). If \( D^2 f(x) \) is negative definite for all \( x \in U \), then \( f \) is strictly concave on \( U \). Conditions for convexity are obtained by replacing "negative" by "positive".

**Remark 2** The conditions for strict concavity in the theorem are only sufficient, not necessary. In other words, if \( D^2 f(x) \) is not negative semidefinite for all \( x \in U \), then \( f \) is not concave; while if \( D^2 f(x) \) is not negative definite for all \( x \in U \), then \( f \) may or may not be strictly concave (see the example below).

**Remark 3** For a matrix \( A \), we often use \( A > 0 \) to denote it is positive definite and \( A \geq 0 \) to denote it is positive semidefinite. \( A < 0 \) and \( A \leq 0 \) can be similarly understood.

We now define and characterize positive/negative (semi)definite matrices.

**Definition 5** An \( n \times n \) matrix \( H \) is **positive definite** iff \( v^T H v > 0 \) for all \( v \neq 0 \) in \( \mathbb{R}^n \); \( H \) is **negative definite** iff \( v^T H v < 0 \) for all \( v \neq 0 \) in \( \mathbb{R}^n \). Replacing the strict inequalities above by weak ones yields the definitions of **positive semidefinite** and **negative semidefinite**.
Remark 4 Usually, positive (negative) definiteness is only defined for a symmetric matrix, so we restrict our discussions on symmetric matrices below. Fortunately, the Hessian is symmetric by Young’s theorem.

Remark 5 The positive definite matrix is an extension of the positive number. To see why? Note that for any positive number $H$, and any real number $v \neq 0$, $v'Hv = v^2H > 0$. Similarly, the positive semidefinite matrix, negative definite matrix, negative semidefinite matrix are extensions of the nonnegative number, negative number and nonpositive number, respectively.

Remark 6 Just like that a function can be neither increasing or decreasing, a matrix can be neither positive (semi)definite or negative (semi)definite, e.g., $H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. (why? Let $v = (1,1)'$ and $(-1,1)'$)

Exercise 18 Show that the diagonal elements of a positive definite matrix must be positive, while the off-diagonal elements need not be.

We now provide some criteria to judge the positive/negative (semi)definiteness. For this purpose, we first define some terms.

Definition 6 For an $n \times n$ matrix $H$, a $k \times k$ submatrix formed by picking out $k$ columns and the same $k$ rows is called a $k$th order principal submatrix of $H$; the determinant of a $k$th order principal submatrix is called a $k$th order principal minor. The $k \times k$ submatrix formed by picking out the first $k$ columns and the first $k$ rows is called a $k$th order leading principal submatrix of $H$; its determinant is called the $k$th order leading principal minor.

Theorem 4 A matrix is positive definite iff its $n$ leading principal minors are all positive. A matrix is negative definite iff its $n$ leading principal minors alternate in sign with the odd order
ones being negative and the even order ones being positive. A matrix is positive semidefinite iff its $2^n - 1$ principal minors are all nonnegative. A matrix is negative semidefinite iff its $2^n - 1$ principal minors alternate in sign so that the odd order ones are nonpositive and the even order ones are nonnegative.

**Example 2**

(i) A linear function $f(x) = a_1x_1 + \cdots + a_nx_n$ is both concave and convex.

(ii) $f(x) = -x^4$ is strictly concave, but its Hessian is not negative definite for all $x \in \mathbb{R}$ since $D^2f(0) = 0$.

(iii) The particular Cobb-Douglas utility function $u(x_1, x_2) = \sqrt{x_1}\sqrt{x_2}$, $(x_1, x_2) \in \mathbb{R}^2_+$, is concave but not strictly concave. First check that it is concave.

$$D^2f(x) = \begin{pmatrix}
\frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_2}}{\sqrt{x_1}} & \frac{1}{2} \left( -\frac{1}{2} \right) \frac{1}{\sqrt{x_1} \sqrt{x_2}} \\
\frac{1}{2} \left( -\frac{1}{2} \right) \frac{1}{\sqrt{x_1} \sqrt{x_2}} & \frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_1}}{\sqrt{x_2}} \\
\end{pmatrix}.$$

Since

$$\frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_2}}{\sqrt{x_1}} \leq 0, \quad \frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_1}}{\sqrt{x_2}} \leq 0$$

and

$$\left( \frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_2}}{\sqrt{x_1}} \right) \left( \frac{1}{2} \left( -\frac{1}{2} \right) \frac{\sqrt{x_1}}{\sqrt{x_2}} \right) - \left( \frac{1}{2} \left( -\frac{1}{2} \right) \frac{1}{\sqrt{x_1} \sqrt{x_2}} \right)^2 = 0$$

for $(x_1, x_2) \in \mathbb{R}^2_+$, $u(x_1, x_2)$ is concave. Now, let $x_2 = x'_2 = 0$, $x_1 \neq x'_1$; then $u(tx_1 + (1-t)x'_1, 0) = 0 = tu(x_1, 0) + (1-t)u(x'_1, 0)$, so $u(x_1, x_2)$ is not strictly concave.

### 2.2 The Uniqueness Theorem

We present two results which indicate the importance of convexity in optimization theory. The first result establishes that in concave optimization problems, all local optima must also be global optima; and, therefore, to find a global optimum in such problems, it always suffices to locate a local optimum. The second result shows that if a strictly concave optimization problem admits a solution, the solution must be unique. We state the two theorem in the context of the mixed constrained maximization problem, i.e.,

$$\max_x f(x)$$

s.t. $g(x) \geq 0$,

$h(x) = 0$,

and suppose the maximizer exists. Recall also that the feasible set $G = \{ x \in \mathbb{R}^n | g(x) \geq 0, h(x) = 0 \}$.

**Theorem 5** If $f$ is concave, and the feasible set $G$ is convex, then

(i) Any local maximum of $f$ is a global maximum of $f$.
(ii) The set \( \arg \max \{ f(x) | x \in G \} \) is convex.\(^1\)

**Theorem 6** If \( f \) is strictly concave, and the feasible set \( G \) is convex, then the maximizer \( x^* \) is unique.

**Proof.** Suppose \( f \) has two maximizers, say, \( x \) and \( x' \); then \( tx + (1-t)x' \in G \), and by the definition of strict concavity, for \( 0 < t < 1 \),

\[
f(tx + (1-t)x') > tf(x) + (1-t)f(x') = f(x) = f(x').
\]

A contradiction. \( \blacksquare \)

**Example 3 (Consumer’s Problem - Revisited)** Does the consumer’s problem

\[
\max_{x_1, x_2} \sqrt{x_1} \sqrt{x_2} \quad \text{s.t.} \quad x_1 + x_2 \leq 1, x_1 \geq 0, x_2 \geq 0
\]

have a solution? Is the solution unique?

**Solution:** The feasible set \( G = \{ x_1 + x_2 \leq 1, x_1 \geq 0, x_2 \geq 0 \} \) is compact (why?) and \( \sqrt{x_1} \sqrt{x_2} \) is continuous, so by the Weierstrass Theorem, there exists a solution.

The solution is unique, \( (x_1^*, x_2^*) = (\frac{1}{2}, \frac{1}{2}) \). But from the discussion above, \( \sqrt{x_1} \sqrt{x_2} \) is not strictly concave for \( (x_1, x_2) \in \mathbb{R}^2_+ \). Actually, even if we restrict \( (x_1, x_2) \in \mathbb{R}^2_{++} \), where \( \mathbb{R}^2_{++} = \{ x | x > 0 \} \), \( \sqrt{x_1} \sqrt{x_2} \) is NOT strictly concave. Check for \( t \in (0, 1), x_1 \neq x_1' \) and/or \( x_2 \neq x_2' \),

\[
\sqrt{tx_1} + (1-t)x_2' \geq t\sqrt{x_1} + (1-t)x_2 \\
\iff (tx_1 + (1-t)x_1') (tx_2 + (1-t)x_2') \geq (t\sqrt{x_1} + (1-t)\sqrt{x_1'})^2 \\
\iff x_1x_2' + x_1'x_2 \geq 2\sqrt{x_1x_2x_1'}x_2' \iff \left( \sqrt{x_1}x_2' - \sqrt{x_1'}x_2 \right)^2 \geq 0
\]

with equality holding when \( x_2/x_1 = x_2'/x_1' \) (what does this mean?).

In summary, the theorem provides only sufficient (but not necessary) conditions. \( \square \)

### 2.3 Quasiconvex Functions (*)

One key question in the uniqueness theorem is how to guarantee that \( G \) is convex. To answer this question, we need to further define a weaker concept than convexity/concavity.

**Definition 7** A function \( f : S \to \mathbb{R} \) defined on a convex set \( S \) is **quasiconvex** if for any \( a \in \mathbb{R} \), the lower contour set \( \{ x | f(x) \leq a \} \) is convex. The negative of a quasiconvex function is said to be **quasiconcave** or its upper contour sets are convex.

\(^1\)Even if \( \arg \max \{ f(x) | x \in G \} = \emptyset \), this is still correct since \( \emptyset \) is convex.
Remark 7 An equivalent definition of quasiconvexity/quasiconcavity is as follows. A function 
\( f : S \rightarrow \mathbb{R} \) defined on a convex set \( S \) is quasiconvex if for all \( x, y \in S \) and \( t \in (0, 1) \) we have
\[
f (tx + (1-t)y) \leq \max \{ f(x), f(y) \}.
\]
A function \( f \) is quasiconcave if
\[
f (tx + (1-t)y) \geq \min \{ f(x), f(y) \}.
\]
Replacing the equality by strict equality, we get the strictly quasiconvex/quasiconcave function.

Remark 8 It is not hard to see that a convex/concave function must be quasiconvex/quasiconcave. The converse is not be true. A concave function can be quasiconvex function; e.g., \( f(x) = \log(x) \) is concave, and it is quasiconvex. The following two figures illustrate a quasiconvex function which is not convex (or concave) and a non-quasiconvex function, respectively.

![A Quasiconvex but Non-Convex Function](image1)

![A Non-QuasiConvex Function](image2)

Remark 9 A convex/concave function must be continuous in the interior of its domain (why the boundary is not included?), while the quasiconvex/quasiconcave function can be discontinuous in the interior of its domain.

Remark 10 A function is both concave and convex iff it is linear (or, more properly, affine), taking the form \( f(x) = a + b'x \) for some constants \( a \) and \( b \). A function that is both quasiconvex and quasiconcave is called quasilinear. Any monotone function on \( \mathbb{R} \) is quasilinear, and any affine function is quasilinear. For a quasilinear function \( f \), \( \{ x | f(x) = c \} \) for any constant \( c \) is convex.

We can now provide some conditions to guarantee \( G \) to be convex. There are many alternative conditions. One popular set of conditions is that \( g_j, j = 1, \ldots, J \), is quasiconcave, and \( h_k, k = 1, \ldots, K \), is quasilinear. For example, in the consumer’s problem above, \( g_1(x) = x_1, g_2(x) = x_2 \) and \( g_3(x) = 1 - x_1 - x_2 \) are all affine, so \( G \) is convex.
(**) In Theorem 6, the strict concavity of $f$ can be replaced by strict quasiconcavity. Also, although for quasiconcave functions, local maxima need not be global maxima, for strict quasiconcave functions this is indeed true. See Theorem 8.12 of Sundaram (1996). (**)

2.4 Sufficient Conditions for Optimization

We now state the Theorem of Kuhn-Tucker under concavity which especially provides some sufficient conditions for optimization.

Theorem 7 (Theorem of Kuhn-Tucker under Concavity) Suppose $f$, $g_j$ and $h_k$, $j = 1, \cdots, J$, $k = 1, \cdots, K$, are all $C^1$ function, $f$ is concave, $g_j$ is quasiconcave, and $h_k$ is quasilinear. If there exists $(\lambda^*, \mu^*)$ such that $(x^*, \lambda^*, \mu^*)$ satisfies the Kuhn-Tucker conditions, then $x^*$ solves the mixed constrained maximization problem.

Remark 11 In practice, check whether $g_j$ is concave, and $h_k$ is affine.

Combining with the necessity of Kuhn-Tucker conditions, we can state the necessary and sufficient conditions for optimization which are omitted here to save space.

3 Second Order Conditions for Optimization (*)

As mentioned in Section 1.2 of Chapter 2, the FOCs cannot determine whether their solutions are (local) maximizers or minimizers. The second order conditions (SOCs) do the job. Also, when the set of critical points is large, the second order conditions (SOCs) may provide further refinements. We will provide both necessary and sufficient SOCs. We first consider the equality-constrained problem and then treat the inequality-constrained problem as an extension.

Theorem 8 In the equality-constrained maximization problem, suppose that $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$ are $C^2$ functions, and the FOCs and NDCQ are satisfied. Define $C(x^*) = \{ v \in \mathbb{R}^n | Dh(x^*)v = 0_K \}$ as the linear constraint set and let the $n \times n$ matrix $D^2L^*$ denote the Hessian matrix of $L$ at $(x^*, \lambda^*)$, i.e., $D^2L^* = D^2_\lambda L(x^*, \lambda^*) = D^2f(x^*) + \sum_{k=1}^{K} \lambda_k^*(D^2h_k(x^*))$.

(i) If $f$ has a local maximum on $G$ at $x^*$, then $\nu'(D^2L^*)v \leq 0$ for all $v \in C(x^*)$.

(ii) If $\nu'(D^2L^*)v < 0$ for all $v \in C(x^*)$ with $v \neq 0$, then $x^*$ is a strict local maximizer of $f$ on $G$.

Exercise 19 State the SOCs for the equality-constrained minimization problem.

Exercise 20 State the SOCs for the unconstrained problem.

Exercise 21 Show that $C(x^*)$ is the tangent space to $\{ h(x) = c \}$ at $x^*$. Hint: Simon and Blume (1994), pp. 459-460.
Result (i) is the second order necessary condition and (ii) is the second order sufficient conditions. These conditions are stated in the form of negative definiteness of a symmetric matrix on the linear constraint set. How to check these conditions? The following theorem answer this question.

We first define some notations. For an \(n \times n\) symmetric matrix \(A\), \(A_l\) denotes the \(l \times l\) submatrix of \(A\) obtained by retaining only the first \(l\) rows and columns of \(A\). For a \(K \times n\) matrix \(B\), \(B_{Kl}\) denotes the \(K \times l\) matrix obtained by retaining only the \(l\) columns of \(B\); when \(K = l\), denote \(B_{Kl}\) as \(B_K\). Given any permutation \(\pi\) of the first \(n\) integers, let \(A^{\pi}\) denote the \(n \times n\) symmetric matrix obtained from \(A\) by applying the permutation \(\pi\) to both its rows and columns and let \(B^{\pi}\) denote the \(K \times n\) matrix obtained by applying the permutation \(\pi\) to only the columns of \(B\). \(A_l^{\pi}\) and \(B_{Kl}^{\pi}\) are the \(A_l\) and \(B_{Kl}\) counterpart of \(A^{\pi}\) and \(B^{\pi}\). Finally, let \(C_l\) be the \((K + l) \times (K + l)\) matrix obtained by "bordering" the submatrix \(A_l\) by the submatrix \(B_{Kl}\) in the following manner:

\[
C_l = \begin{pmatrix}
0_K & B_{Kl} \\
B_{Kl}' & A_l
\end{pmatrix}.
\]

Denote by \(C_l^{\pi}\) the obtained similarly when \(A\) is replaced by \(A^{\pi}\) and \(B\) by \(B^{\pi}\). For any matrix \(A\), \(|A|\) denotes \(A\)'s determinant.

**Theorem 9** Let \(A\) be a symmetric \(n \times n\) matrix, and \(B\) a \(K \times n\) matrix such that \(|B_K| \neq 0\). Define the bordered matrices \(C_l\) as described above. Then,

(i) \(x'Ax \geq 0\) for every \(x\) such that \(Bx = 0\) iff for all permutations \(\pi\) of the first \(n\) integers, and for all \(r \in \{K + 1, \cdots, n\}\), we have \((-1)^K |C_l^{\pi}| \geq 0\).

(ii) \(x'Ax \leq 0\) for every \(x\) such that \(Bx = 0\) iff for all permutations \(\pi\) of the first \(n\) integers, and for all \(r \in \{K + 1, \cdots, n\}\), we have \((-1)^r |C_r^{\pi}| \geq 0\).

(iii) \(x'Ax > 0\) for every \(x\) such that \(Bx = 0\) iff for all \(r \in \{K + 1, \cdots, n\}\), we have \((-1)^K |C_r| > 0\).

(ii) \(x'Ax < 0\) for every \(x\) such that \(Bx = 0\) iff for all \(r \in \{K + 1, \cdots, n\}\), we have \((-1)^r |C_r| > 0\).

When \(A = D^2\mathcal{L}^*\) and \(B = Dh(x^*)\), the matrices \(C_l\) are called "bordered Hessians". We assume \(|B_K| \neq 0\) because rank(\(B\)) = \(K\) by the NDCQ and it is without loss of generality to assume the first \(K\) columns of \(B\) to be linearly independent.

**Exercise 22** Show that the criteria in the above theorem degenerate to those in Theorem 4 when \(K = 0\).

**Exercise 23** Show that when \(A = D^2\mathcal{L}^*\) and \(B = Dh(x^*)\), \(C_n = D_{\lambda, x}^2\mathcal{L}(x^*, \lambda^*)\).

In the inequality-constrained problem, we can replace \(h\) by the union of \(h\) and the binding constraints at \(x^*\) in \(g\). Since for the unbinding constraints the corresponding \(\lambda^*\)'s are zero, the inequality-constrained problem reduces to an equality-constrained problem.
Using the bordered Hessians to check $x^*$ is a local maximizer or a local minimizer may be burdensome in practice. As an easy (although less general) alternative, we can employ the concavity of the objective function $f$ to draw the conclusion. Specifically, if $f$ is strictly concave at $x^*$, i.e., $D^2 f(x^*) < 0$, then $x^*$ is a strict local maximizer; if $f$ is strictly convex at $x^*$, i.e., $D^2 f(x^*) > 0$, then $x^*$ is a strict local minimizer. This is actually the second order sufficient conditions for the unconstrained problem in Exercise 20.

The bordered Hessians can also be used to judge whether a function is quasiconcave/quasiconvex or not. First define

$$C_l(x) = \begin{pmatrix} 0 & Df(x)_l \\ Df(x)_l & D^2 f(x)_l \end{pmatrix}, l = 1, \ldots, n,$$

where $Df(x)_l = \left( \frac{\partial f}{\partial x_1} (x), \ldots, \frac{\partial f}{\partial x_n} (x) \right)$ and $D^2 f(x)_l$ is similarly defined as $A_l$.

**Theorem 10** Let $f \in C^2(U)$, where $U \subset \mathbb{R}^n$ is open and convex. If $f$ is quasiconcave, then $(-1)^r |C_r(x)| \geq 0$, $r = 1, \ldots, n$; if $(-1)^r |C_r(x)| > 0$, $r = 1, \ldots, n$, then $f$ is strictly quasiconcave. Conditions for quasiconvexity are obtained by omitting $(-1)^r$.

**Remark 12** In Theorem 3 a weak inequality (viz., the negative semidefiniteness of $D^2 f(x)$) was both necessary and sufficient for concavity. In the current result, the weak inequality is necessary but insufficient for quasiconcavity; see Example 8.10 of Sundaram (1996). Nevertheless, to check negative semidefiniteness, we need to check all principal submatrices rather than only leading principal submatrices as $C_l(x)$.

## 4 Appendix A: Support and Separation (*)

Another set of mathematical results closely connected to the notion of convexity is so-called separation and support theorems. These theorems are frequently used in economics to obtain a price system that leads consumers and producers to choose Pareto-efficient allocation. That is, given the prices, producers are maximizing profits, and given those profits as income, consumers are maximizing utility subject to their budget constraints. This is so-called the **second fundamental theorem of welfare economics**.

### 4.1 Hyperplanes

The concept of hyperplane in $\mathbb{R}^n$ is a straightforward generalization of the notion of a line in $\mathbb{R}^2$ and of a plane in $\mathbb{R}^3$. A line in $\mathbb{R}^2$ can be described by an equation

$$p_1 x_1 + p_2 x_2 = \alpha,$$

where $p = (p_1, p_2)'$ is some non-zero vector and $\alpha$ is some scalar. A plane in $\mathbb{R}^3$ can be described by an equation

$$p_1 x_1 + p_2 x_2 + p_3 x_3 = \alpha,$$
where \( \mathbf{p} = (p_1, p_2, p_3)' \) is some non-zero vector and \( \alpha \) is some scalar. Similarly, a hyperplane in \( \mathbb{R}^n \) can be described by an equation
\[
\sum_{i=1}^{n} p_i x_i = \alpha
\]
where \( \mathbf{p} = (p_1, p_2, \ldots, p_n)' \) is some non-zero vector in \( \mathbb{R}^n \) and \( \alpha \) is some scalar. It can be written in more concise way using scalar (aka inner, dot) product notation.

**Definition 8** A hyperplane is the set
\[
H(\mathbf{p}, \alpha) = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{x} = \alpha \}
\]
where \( \mathbf{p} \in \mathbb{R}^n \) is a non-zero vector and \( \alpha \) is a scalar. The vector \( \mathbf{p} \) is called the normal to the hyperplane \( H \).

Suppose that there are two points \( \mathbf{x}^*, \mathbf{y}^* \in H(\mathbf{p}, \alpha) \). Then by definition \( \mathbf{p} \cdot \mathbf{x}^* = \alpha \) and \( \mathbf{p} \cdot \mathbf{y}^* = \alpha \). Hence \( \mathbf{p} \cdot (\mathbf{x}^* - \mathbf{y}^*) = 0 \). In other words, vector \( \mathbf{p} \) is orthogonal to the line segment \( (\mathbf{x}^* - \mathbf{y}^*) \). Since we started by picking arbitrary points in \( H(\mathbf{p}, \alpha) \), we have that \( \mathbf{p} \) is orthogonal to any line segment in \( H(\mathbf{p}, \alpha) \), or that \( \mathbf{p} \) is orthogonal to \( H(\mathbf{p}, \alpha) \).

Given a hyperplane \( H \subset \mathbb{R}^n \), points in \( \mathbb{R}^n \) can be classified according to their positions relative to hyperplane. The (closed) half-space determined by the hyperplane \( H(\mathbf{p}, \alpha) \) is either the set of points ‘below’ \( H \) or the set of points ‘above’ \( H \), i.e., either the set \( \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{x} \leq \alpha \} \) or the set \( \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{x} \geq \alpha \} \). Open half-spaces are defined by strict inequalities.

**Exercise 24** Prove that a closed half-space is closed and open half-space is open.

The straightforward economic example of a half-space is a budget set \( \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{x} \leq \alpha \} \) of a consumer with income \( \alpha \) facing the vector of prices \( \mathbf{p} \). (It was rather neat to call the normal vector \( \mathbf{p} \), wasn’t it?). By the way, hyperplanes and half-spaces are convex sets.

**Exercise 25** Prove that any half-space, open or closed, and any hyperplane in \( \mathbb{R}^n \) is convex.

### 4.2 Support Functions

In this section, we give a description of what is called a dual structure. Consider the set of all closed convex subsets of \( \mathbb{R}^n \). We will show that to each such set \( S \) we can associate an extended-real valued function \( \mu_S : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\} \), that is a function that maps each vector in \( \mathbb{R}^n \) to either a real number or to \(-\infty\). Not all such functions can be arrived at in this way. In fact, we shall show that any such function must be concave and homogeneous of degree 1. But once we restrict attention to functions that can be arrived at as a “support function” for some such closed convex set we have another set of objects that we can analyze and perhaps make useful arguments about the original sets in which we were interested.

In fact, we shall define the function \( \mu_S \) for any subset of \( \mathbb{R}^n \), not just the closed and convex ones. However, if the original set \( S \) is not a closed convex one we shall lose some information about
\( S \) in going to \( \mu_S \). In particular, \( \mu_S \) only depends on the closed convex hull of \( S \), that is, if two sets have the same closed convex hull they will lead to the same function \( \mu_S \).

We define \( \mu_S : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\} \) as

\[
\mu_S(p) = \inf \{ p \cdot x \mid x \in S \},
\]

where \( \inf \) denotes the infimum or greatest lower bound. It is a property of the real numbers that any set of real numbers has an infimum. Thus \( \mu_S(p) \) is well defined for any set \( S \). If the minimum exists, for example if the set \( S \) is compact, then the infimum is the minimum. In other cases the minimum may not exist. To take a simple one dimensional example, suppose that the set \( S \) was the subset of \( \mathbb{R} \) consisting of the numbers \( 1/n \) for \( n = 1, \ldots \) and that \( p = 2 \). Then clearly \( p \cdot x = px \) does not have a minimum on the set \( S \). However, 0 is less than \( px = 2x \) for any value of \( x \) in \( S \) but for any number \( a \) greater than 0 there is a value of \( x \) in \( S \) such that \( px < a \). Thus 0 is in this case the infimum of the set \( \{ p \cdot x \mid x \in S \} \).

Recall that we have not assumed that \( S \) is convex. However, if we do assume that \( S \) is both convex and closed then the function \( \mu_S \) contains all the information needed to reconstruct \( S \).

Given any extended-real valued function \( \mu : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\} \), let us define the set \( S_\mu \) as

\[
S_\mu = \{ x \in \mathbb{R}^n \mid p \cdot x \geq \mu(p) \text{ for every } p \in \mathbb{R}^n \}.
\]

That is, for each \( p \) such that \( \mu(p) > -\infty \) we define the closed half space

\[
\{ x \in \mathbb{R}^n \mid p \cdot x \geq \mu(p) \}.
\]

Notice that if \( \mu(p) = -\infty \), then \( p \cdot x \geq \mu(p) \) for any \( x \) and so the above set will be \( \mathbb{R}^n \) rather than a half space and that for this \( p \) the requirement that \( p \cdot x \geq \mu(p) \) puts no restrictions on the set \( S_\mu \). The set \( S_\mu \) is the intersection of all these closed half spaces. Since the intersection of convex sets is convex and the intersection of closed sets is closed, the set \( S_\mu \) is, for any function \( \mu \), a closed convex set.

Suppose that we start with a set \( S \), define \( \mu_S \) as above and then use \( \mu_S \) to define the set \( S_{\mu_S} \). If the set \( S \) was a closed convex set, then \( S_{\mu_S} \) will be exactly equal to \( S \). Since we have seen that \( S_{\mu_S} \) is a closed convex set, it must be that if \( S \) is not a closed convex set it will not be equal to \( S_{\mu_S} \). However, \( S \) will always be a subset of \( S_{\mu_S} \), and indeed \( S_{\mu_S} \) will be the smallest closed convex set such that \( S \) is a subset, that is, \( S_{\mu_S} \) is the closed convex hull of \( S \).

So we see that the properties of \( \mu_S \) do not depend on the set \( S \) being closed or convex. Whether or not \( S \) is closed or convex \( \mu_S \) will depend only on the closed convex hull of \( S \). We obtain a similar result about the function \( \mu \) and the process of going from \( \mu \) to \( S_\mu \) and then to \( \mu_{S_\mu} \).

### 4.3 Separation

We now consider the notion of ‘separating’ two sets by a hyperplane.
Definition 9 A hyperplane $H$ separates sets $A$ and $B$ if $A$ is contained in one closed half-space and $B$ is contained in the other. A hyperplane $H$ strictly separates sets $A$ and $B$ if $A$ is contained in one open half-space and $B$ is contained in the other.

It is clear that strict separation requires the two sets to be disjoint. For example, consider two (externally) tangent circles in a plane. Their common tangent line separates them but does not separate them strictly. On the other hand, although it is necessary for two sets be disjoint in order to strictly separate them, this condition is not sufficient, even for closed convex sets. Let $A = \{x \in \mathbb{R}^2 \mid x_1 > 0 \text{ and } x_1 x_2 \geq 1\}$ and $B = \{x \in \mathbb{R}^2 \mid x_1 \geq 0 \text{ and } x_2 = 0\}$ then $A$ and $B$ are disjoint closed convex sets but they cannot be strictly separated by a hyperplane (line in $\mathbb{R}^2$). Thus the problem of the existence of separating hyperplane is more involved than it may appear to be at first.

We start with separation of a set and a point.

Theorem 11 Let $S \subseteq \mathbb{R}^n$ be a convex set and $x_0 \notin S$ be a point. Then $S$ and $x_0$ can be separated. If $S$ is closed then $S$ and $x$ can be strictly separated.

Idea of proof. Proof proceeds in two steps. The first step establishes the existence a point $a$ in the closure of $S$ which is the closest to $x_0$. The second step constructs the separating hyperplane using the point $a$.

STEP 1. There exists a point $x \in \bar{S}$ (closure of $S$) such that $d(x_0, a) \leq d(x, a)$ for all $x \in \bar{S}$, and $d(x_0, x) \geq 0$.

Let $B(x_0)$ be closed ball with center at $x_0$ that intersects the closure of $S$. Let $A = B(x_0) \cap \bar{S} \neq \emptyset$. The set $A$ is nonempty, closed and bounded (hence compact). According to Weierstrass’s theorem, the continuous distance function $d(x_0, x)$ achieves its minimum in $A$. That is, there exists $a \in A$ such that $d(x_0, a) \leq d(x, a)$ for all $x \in \bar{S}$. Note that $d(x_0, a) \geq 0$.

STEP 2. There exists a hyperplane $H(p, \alpha) = \{x \in \mathbb{R}^n \mid p \cdot x = \alpha\}$ such that $p \cdot x \geq \alpha$ for all $x \in \bar{S}$ and $p \cdot x_0 \leq \alpha$.

Construct a hyperplane which goes through the point $a \in \bar{S}$ and has normal $p = a - x_0$. The proof that this hyperplane is the separating one is done by contradiction. Suppose there exists a point $y \in \bar{S}$ which is strictly on the same side of $H$ as $x_0$. Consider the point $y' \in [a, y]$ such that the vector $y' - x_0$ is orthogonal to $y - a$. Since $d(x_0, y) \geq d(x_0, a)$, the point $y'$ is between $a$ and $y$. Thus, $y \in \bar{S}$ and $d(x_0, y') \leq d(x_0, a)$ which contradicts the choice of $a$. When $S = \bar{S}$, that is $S$ is closed, the separation can be made strict by choosing a point strictly in between $a$ and $x_0$ instead of $a$. This is always possible because $d(x_0, a) > 0$ when $S$ is closed.

Theorem 11 is very useful because separation of a pair of sets can be always reduced to separation of a set and a point.

Lemma 2 Let $A$ and $B$ be a non-empty sets. $A$ and $B$ can be separated (strictly separated) iff $A - B$ and 0 can be separated (strictly separated).
Theorem 12 (Minkowski, 1911) Let \( A \) and \( B \) be non-empty convex sets with \( A \cap B = \emptyset \). Then \( A \) and \( B \) can be separated. If \( A \) is compact and \( B \) is closed then \( A \) and \( B \) can be strictly separated.

Proof. If \( A \) and \( B \) are convex, then \( A - B \) is convex. If \( A \) is compact and \( B \) is closed, then \( A - B \) is closed. And \( 0 \notin A - B \) iff \( A \cap B = \emptyset \). 

An extension of the separation theorem to the infinite-dimensional space is the famous Hahn-Banach separation theorem. To state the theorem, we first define a locally convex vector space.

Definition 10 A topological vector space \( V \) is said to be locally convex, if every point has a fundamental system of convex open neighborhoods. This means that for every \( v \in V \) and every neighborhood \( N \) of \( x \), there exists a convex open set \( D \), with \( v \in D \subseteq N \).

Remark 13 Every normed space is a Hausdorff locally convex space.

Theorem 13 (Hahn-Banach Separation Theorem) Let \( V \) be a topological vector space. If \( A, B \) are convex, non-empty disjoint subsets of \( V \), then:

(i) If \( A \) is open, then there exists a continuous linear map \( \lambda : V \to \mathbb{R} \) and \( t \in \mathbb{R} \) such that \( \lambda(a) < t \leq \lambda(b) \) for all \( a \in A, b \in B \).

(ii) If \( V \) is locally convex, \( A \) is compact, and \( B \) closed, then there exists a continuous linear map \( \lambda : V \to \mathbb{R} \) and \( s, t \in \mathbb{R} \) such that \( \lambda(a) < t < s < \lambda(b) \) for all \( a \in A, b \in B \).

4.4 Support

Closely (not in the topological sense) related to the notion of a separating hyperplane is the notion of supporting hyperplane.

Definition 11 The hyperplane \( H \) supports the set \( S \) at the point \( x_0 \in S \) if \( x_0 \in H \) and \( S \) is a subset of one of the closed half-spaces determined by \( H \).

A convex set can be supported at any of its boundary points, this is the immediate consequence of Theorem \[12\]. To prove it, consider the sets \( A \) and \( B = \{x_0\} \), where \( x_0 \) is a boundary point of \( A \).

Theorem 14 Let \( S \subseteq \mathbb{R}^n \) be a convex set with nonempty interior and \( x_0 \in S \) be its boundary point. Then there exist a supporting hyperplane for \( S \) at \( x_0 \).

Note that if the boundary of a convex set is smooth (‘differentiable’) at the given point \( x_0 \) then the supporting hyperplane is unique and is just the tangent hyperplane. If, however, the boundary is not smooth then there can be many supporting hyperplanes passing through the given point. It is important to note that conceptually the supporting theorems are connected to calculus. But, the supporting theorems are more powerful (don’t require smoothness), more direct, and more set-theoretic.

Certain points on the boundary of a convex set carry a lot of information about the set.
Definition 12 A point \( x \) of a convex set \( S \) is an **extreme point** of \( S \) if \( x \) is not an interior point of any line segment in \( S \).

The extreme points of a closed ball and of a closed cube in \( \mathbb{R}^3 \) are its boundary points and its eight vertices, respectively. A half-space has no extreme points even if it is closed.

An interesting property of extreme points is that an extreme point can be deleted from the set without destroying convexity of the set. That is, a point \( x \) in a convex set \( S \) is an extreme point iff the set \( S\setminus \{x\} \) is convex.

The next Theorem is a finite-dimensional version of a quite general and powerful result by M.G. Krein and D.P. Milman.

**Theorem 15 (Krein & Milman, 1940)** Let \( S \subseteq \mathbb{R}^n \) be convex and compact. Then \( S \) is the convex hull of its extreme points.

5 Appendix B: Fixed Point Theorems (*)

We list a few most popular fixed point theorems below, which are useful in proving the existence of equilibria or optimal solutions. More general discussions are referred to Franklin (1980), Border (1985) and Zeidler (1986). We list these fixed point theorem here because some of them are related to convexity.

The first fixed point theorem is the contraction mapping theorem developed in Banach (1922).

**Theorem 16 (Contraction Mapping Theorem)** Let \( (M, d) \) be a complete metric space. Let \( f : M \to M \) be a contraction. That is, for any \( x, y \in M \) we have \( d(f(x), f(y)) \leq \theta d(x, y) \) where \( 0 \leq \theta < 1 \). Then \( f \) has a unique fixed point in \( M \), i.e., there is a unique point \( x^* \in M \) such that \( x^* = f(x^*) \).

The second fixed point theorem is the famous Brouwer’s fixed point theorem. In economics, Brouwer’s fixed-point theorem and its extension, the Kakutani fixed-point theorem, play a central role in the proof of existence of general equilibrium in market economies as developed in the 1950s by economics Nobel prize winners Kenneth Arrow and Gérard Debreu.

**Theorem 17 (Brouwer, 1910)** Let \( M \subseteq \mathbb{R}^n \) be a compact convex set. Let \( f : M \to M \) be a continuous function. Then there is some \( x^* \in M \) such that \( x^* = f(x^*) \).

**Proof.** Franklin (1980) contains three different proofs. ■

The Kakutani fixed-point theorem involves correspondences or multivalued functions which are less standard in the mathematics literature, though are of central importance in economics and game theory. Some materials on this topic are covered in Berge (1963); more details can be found in Klein and Thompson (1984).

To state the Kakutani fixed-point theorem, we first define upper hemicontinuity.


Definition 13 (Correspondence) A correspondence $F$ between the sets $X$ and $Y$, written $F : X \to Y$, is a function from the set $X$ to the set of all subsets of $Y$ (written $2^Y$) such that for all $x \in X$, $F(x) \neq \emptyset$.

Intuitively, a correspondence $F : X \to Y$ is said to be upper hemicontinuous (uhc) at the point $a$ if for any open neighborhood $V$ of $F(a)$ there exists a neighborhood $U$ of $a$ such that for all $x$ in $U$, $F(x)$ is a subset of $V$.

Definition 14 (Upper Hemicontinuity) For a correspondence $F : X \to Y$ with closed values (i.e., $F(x)$ is closed in $Y$ for all $x \in X$), $F : X \to Y$ is upper hemicontinuous at $x \in X$ if $\forall x_n \in X$, $\forall y \in Y$ and $\forall y_n \in F(x_n)$,

$$\lim_{n \to \infty} x_n \to x, \lim_{n \to \infty} y_n \to y \implies y \in F(x)$$

If $F$ is compact-valued (i.e. $F(x)$ is compact for all $x \in X$) the converse is also true.

When we confine our attention to the case that the space $Y$ is compact, there is an easier way to make the definition.

Theorem 18 (Closed Graph Theorem) Let $X$ be a topological space and $Y$ be a compact topological space. Let $F : X \to Y$ be a closed valued correspondence. Then $F$ is upper hemicontinuous iff the graph of $F$, $G_F = \{(x, y) | x \in X, y = T(x)\}$, is closed in $X \times Y$.

Parallelly, we can define lower hemicontinuity. Intuitively, a correspondence $F : X \to Y$ is said to be lower hemicontinuous (lhc) at the point $a$ if for any open set $V$ intersecting $F(a)$ there exists a neighborhood $U$ of $a$ such that $F(x)$ intersects $V$ for all $x$ in $U$. (Here $V$ intersects $S$ means nonempty intersection $V \cap S \neq \emptyset$).

Definition 15 (Lower Hemicontinuity) A correspondence $F : X \to Y$ is lower hemicontinuous at $x \in X$ if $\forall x_n \in X$ with $\lim_{n \to \infty} x_n \to x$, $\forall y \in F(x)$, $\exists$ a subsequence $x_{n_k}$ of $x_n$ and $y_k \in F(x_{n_k})$ such that $y_k \to y$.

Theorem 19 (Open Graph Theorem) If $F : X \to Y$ has open graph $G_F$, then it is lower hemicontinuous.

Figure ?? illustrates various cases that are and are not upper and/or lower hemicontinuous. Intuitively, if $F$ is uhc, then it can explode in the limit, but cannot implode; if $F$ is lhc, then it can implode in the limit, but cannot explode. If a correspondence is both upper hemicontinuous and lower hemicontinuous, it is said to be continuous.

Finally, we state the Kakutani fixed-point theorem.

Theorem 20 (Kakutani, 1941) Let $M \subset \mathbb{R}^n$ be a compact convex set. Let $F : M \to M$ be an upper hemicontinuous convex valued correspondence. Then there is some $x^* \in M$ such that $x^* = F(x^*)$. 20
**Proof.** See Franklin (1980) and Kakutani (1941).

The following Schauder’s Theorem extends Brouwer’s fixed point theorem to the more general Banach space.

**Theorem 21 (Schauder, 1930)** Let $M$ be a nonempty convex subset of a Banach space $X$. Let $N$ be a compact subset of $M$. Let $f : M \to N$ be a continuous function. Then there is some $x^* \in M$ such that $x^* = f(x^*)$.

**Proof.** See Franklin (1980).

The following is an immediate corollary to Schauder’s Theorem.

**Theorem 22 (Tychono¤, 1935)** Let $M$ be a nonempty compact convex subset of a Banach space $X$. Let $f : M \to M$ be a continuous function. Then there is some $x^* \in M$ such that $x^* = f(x^*)$.

The following Fan-Glicksberg Theorem extends the Kakutani fixed-point theorem to the more general Hausdorff space.

**Theorem 23 (Fan, 1952; Glicksberg, 1952)** Let $M$ be a nonempty compact convex subset of a convex Hausdorff topological vector space. Let $F : M \to M$ be an upper hemicontinuous convex valued correspondence. Then there is some $x^* \in M$ such that $x^* = F(x^*)$. 

---

**Figure 4: Upper and Lower Hemicontinuity**

![Graphical representation of upper and lower hemicontinuity](image-url)