Michigan State EC812A Organized Lecture Notes

Willy Chen and Lauren Koleda (From Jon Eguia Lectures)*

Fall 2022

Contents

*Last Updated: October 1, 2024

1 Set Theory

1.1 Set Operations

Naively, we can define a **set**[1] to be a collection of objects[2]. We define **elements**[3] to be the objects that are in a set. Moreover, we know that any element is either "in" ($\in$) or "not in" ($\notin$) a given set. For example, object a in set A is written as "$a \in A$".

Definition: Z is said to be **Contained** (denoted as $Z \subseteq X$) in a set X if $\forall z \in Z$, $z \in X$.

Definition: Z is said to be **Properly Contained** (denoted as $Z \subset X$) in a set X if $\forall z \in Z$, $z \in X$ AND $\exists x \in X$, $x \notin Z$.

Definition: Sets A and B are said to be **Equal** (denoted as $A = B$) if $A \subseteq B \wedge B \subseteq A$.

Definition (Set Exclusion): Given sets A and B, $B \setminus A = \{x \in B \mid x \notin A\}$[4].

Definition (Union): Given sets A and B, the **Union** of A and B is

$$A \cup B \equiv \{x \mid x \in A \vee x \in B\}$$

Definition (Intersect): Given sets A and B, the **Intersect** of A and B is

$$A \cap B \equiv \{x \mid x \in A \wedge x \in B\}$$

Definition (Singleton): A set A is a singleton if "$x \in A \wedge y \in A \iff x = y$".

Definition (Power Set): The **Power Set** of the set X (denoted as 2^X) is defined as

$$2^X \equiv \{A | A \subseteq X\}$$

Definition (Cartesian Product): Let $X_1, X_2, \ldots, X_k$ be non-empty sets. The **Cartesian Product** of these sets is the collection of ordered tuples containing elements from these sets.

[1] Conventionally denoted as capital letters like A, B, C, etc.
[2] For a more formal definition, consult the Zermelo-Fraenkel Axioms
[3] Conventionally denoted as lowercase letters like a, b, c, etc.
[4] Read as "B minus A"

Formally, the *Cartesian Product* of $X_1, X_2, \ldots, X_k$ is written as:

$$X_1 \times X_2 \times \ldots \times X_k \equiv \prod_{i=1}^{k} X_i \equiv \{(x_1, x_2, \ldots, x_k) \mid x_1 \in X_1,\, x_2 \in X_2, \ldots,\, x_k \in X_k\}$$

Definition (Binary Relations): Let X, Y be arbitrary sets. The **Binary Relation** B from X (**Domain**) to Y (**Co-Domain**) is a subset of $X \times Y$. i.e., $B \subseteq X \times Y$ (also written as $x_a B y_a$).

Definition (Correspondence): A **Correspondence** g from X to Y is a binary relation from X to Y such that $\forall x \in X, \exists y \in Y, (x, y) \in g$, $g \subseteq X \times Y \equiv g : X \rightrightarrows Y$. $g(x) \equiv \{y \in Y \mid (x, y) \in g\}$ is said to be the image of the correspondence.

Definition (Function): A binary relation f from X to Y is called a **Function** if it is a correspondence from X to Y such that $\forall x \in X$, $(x, y) \in f$, $(x, z) \in f \Rightarrow y = z$. $g : X \rightrightarrows Y \equiv g : X \to 2^Y \setminus \varnothing$.

Definition: The **Range** of the correspondence $g : X \rightrightarrows Y$ is denoted as $Ran(g) \equiv \{y \in Y \mid \exists x \in X \; s.t. \, (x, y) \in g\}$. Note that range is always either a subset of the Co-Domain.

Definition (Surjection): A function $f : X \to Y$ is said to be **Surjective (on to)** if $Ran(f) = y$.

Definition (Injection): A function $f : X \to Y$ is said to be **Injective (one-to-one)** if $f(x_1) = f(x_2) \iff x_1 = x_2$.

Definition: A function $f : X \to Y$ is said to be **Bijective** if it is both surjective and injective. A function f is **Invertible** if and only if it is bijective.

Definition: A set X is said to be **Closed under *f*** if $f(X) \subseteq X$

1.2 Constructing Numbers

Naturally, we use the empty set $\varnothing$ to denote the collection of nothing-ness. Think of this as an empty box. Then we can also think of a set of the empty set $\{\varnothing\}$ (think of this as a box that has an empty box inside). For generality, we assign the symbol 1 to the box that contains exactly one empty box. i.e., $1 \equiv \{\varnothing\}$.

For any set A, define the successor function as $S(A) \equiv A \cup \{A\}$. We can then define the ordered successors of 1 as:

$$\begin{aligned}
2 \equiv & S(1) = \{\varnothing\} \cup \{\{\varnothing\}\} = \{ \underbrace{\varnothing}_{\text{First object}}, \underbrace{\{\varnothing\}}_{\text{Second object}} \} \\
3 \equiv & S(2) = 2 \cup \{2\} = \{\varnothing\} \cup \{\{\varnothing\}\} \cup \{\varnothing\} \cup \{\{\varnothing\}\}\} \\
& = \{ \underbrace{\varnothing}_{\text{First object}}, \underbrace{\{\varnothing\}}_{\text{Second object}}, \underbrace{\{\varnothing, \{\varnothing\}\}}_{\text{Third object}}\} \\
4 \equiv & S(3) \\
& \vdots
\end{aligned}$$

One can conclude that our definition is that each number n is the set with n distinct objects.

Definition: $\mathbb{N}$ is the smallest collection of sets that contains 1 and is closed under the successor function $S(\cdot)$[5].

Definition: X is said to be **Countable** if there exists a **Bijective** function $f : X \to \mathbb{N}$ such that $f(X) = N \subseteq \mathbb{N}$.

We can define two relations $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ and $\geq$ that follow the following properties:

1. **Commutative** $+$: $\forall\, m, n, k \in \mathbb{N}$, $((m, n), k) \in + \Rightarrow\ ((n, m), k) \in +$
2. **Order relation** $(>, =, <)$: $\geq \equiv \{(n, m) \in \mathbb{N}^2 \mid n \geq m\}$

Definition: Let $\leq \subseteq X^2$, then $(X, \leq)$ is **Linear Order** if:

1. **Reflexivity:** $x \leq x, \forall x \in X$ (Each element is related to itself)
2. **Anti-Symmetry:** $(x \leq y) \wedge (y \leq x) \iff x = y$
3. **Transitivity:** $(x \leq y) \wedge (y \leq z) \iff x \leq z$
4. **Completeness** $\forall\, x, y \in X$, either $x \leq y$ or $y \leq x$

So far, we have defined the set of natural numbers with an additive operation and an order relation $(\mathbb{N}, +, \geq)$, but based on our experience with mathematics at this point, that's not nice enough, is it?

[5]A set X is closed under the operation $\circ$ if $\forall a, b \in X$, $a \circ b \in X$

Consider a new set $\mathbb{Z}$ that consists of $\mathbb{N}$ as well as something else. To make it reasonably nice, we need to make sure that we can get from elements in $\mathbb{Z}$ to another element in $\mathbb{Z}$. i.e., we need the additive inverses of $\mathbb{N}$ to be in the set $\mathbb{Z}$ as well as a basic **identity** element in $\mathbb{Z}$. So we have $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \{-x \mid x \in \mathbb{N}\}$

Definition: $(X, *)$ is a **Abelian/Commutative Group**[6] if the following properties hold

(i) X is closed under $*$: $\forall x, y \in X$, $x * y \in X$

(ii) $*$ is associative: $\forall x, y, z \in X$, $(x * y) * z = x * (y * z)$

(iii) Identity element u: $\exists u \in X$ such that $\forall x \in X$, $x * u = u * x = x$

(iv) Inverse element y: $\forall x \in X$, $\exists y \in X$ such that $x * y = u$

(v) $*$ is commutative: $\forall x, y \in X$, $x * y = y * x$

One can check that $(\mathbb{Z}, +)$ as we have defined is an abelian group. Moreover, since $\mathbb{Z}$ is defined from $\mathbb{N}$ and its identity and inverse elements, we know that the order relations $\geq$ from before also hold. At this point, you are probably thinking that this is nice, but it can be nicer, and you would be right!

Notice that multiplication $(\times)$ is implied from addition $(+)$, but it is trivial that we cannot necessarily find the inverses of multiplication.

Definition: Let X be a set that $*, \cdot$ two operations on the set X. $(X, *, \cdot)$ is said to be a **Field** if

(i) $(X, *)$ is an abelian group

(ii) $(X, \cdot)$ is closed under $\cdot$

(iii) $\cdot$ is commutative

(iv) $\cdot$ is associative

(v) $(X, \cdot)$ has an identity element

(vi) Any element of X except the identity element under $*$ has an inverse element under $\cdot$

[6] All groups satisfy (1)-(4)

Naturally, define $Q \equiv \{(m,n) \in \mathbb{Z}^2 \mid n \neq 0\}$. From the definition above, we know that $(Q,+,\times)$ is a field, moreover, $\geq$ is still preserved.
We can define the absolute value function $|\cdot| : \mathbb{Q} \to \mathbb{Q}_+$ as

$$|x| = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$$

Now $\mathbb{Q}$ is really nice, but over the course of history, mathematicians realized that it is not nice enough as they started to find little holes between elements of $\mathbb{Q}$. Interestingly, there are actually more holes than non-holes, but we will get to that later.

Definition: Take $(X, \geq)$ to be a linearly ordered set and $S \subset X$. Then $z \in X$ is called a **Supremum** or *least upper bound* of S (denoted as $z = sup(S)$ if:

(i) $\forall\, s \in S,\ s \leq z$

(ii) $\forall\, w \in X, s \in S, z \leq w \Rightarrow s \leq w$

Definition: Similarly, $z \in X$ is called a **Infimum** or *greatest lower bound* of S (denoted as $z = \inf(S)$ if:

(i) $\forall\, s \in S,\ s \geq z$

(ii) $\forall\, w \in X, s \in S, z \geq w \Rightarrow s \geq w$

Definition: Take two sets X, Y and $f : Y \to X$, let $(X, \leq)$ be linearly ordered, then

$$\sup(f(y)) \equiv \sup(f(Y))$$
$$\inf(f(y)) \equiv \inf(f(Y))$$

Remark: Readers should note that the supremum and infimum need not be in the set itself.

Definition: An infinite sequence $(x_n)_{n=1}^{\infty}$ on X is an element of X^∞ that is an ordered tuple such that $(x_k) = (x_1, x_2, \ldots, x_k)$. x_k is called the k^{th} term of the sequence.

Definition: $(x_n)_{n=1}^{\infty}$ on $\mathbb{Q}$ is a **Cauchy sequence** if $\forall \varepsilon \in \mathbb{Q}, \exists m_\varepsilon \in \mathbb{N}\ s.t. \forall n > m_\varepsilon, n \in \mathbb{N}$, $|x_n - x_{m_\varepsilon}| < \varepsilon$

We can think of the holes between numbers in $\mathbb{Q}$ as the limits of the Cauchy sequences from $\mathbb{Q}$ that converge outside of $\mathbb{Q}$, and filling those holes gives us $\mathbb{R}$.

Definition: $\mathbb{R}$ is the set that satisfies:

(i) $(R, +, \cdot)$ is a field

(ii) $\mathbb{R}$ has a complete ordering $(\leq)$ that is compatible with $+, \cdot$

(a) $\forall\, x, y, z \in \mathbb{R},\ x \leq y \Rightarrow x + z \leq y + z$

(b) $\forall\, x, y, z \in \mathbb{R},\ (x \leq y) \wedge (z \geq 0) \Rightarrow x \cdot z \leq y \cdot z$

(iii) $\mathbb{R}$ needs to be complete[7]: Take $L, H \subset \mathbb{R}$ such that $\forall l \in L$, $\forall h \in H$, and $l \leq h$, then $\exists x \in \mathbb{R}$ such that $l \leq x \leq h$

An alternative way to state completeness is—**Definition:** Every non-decreasing/*non-increasing* sequence $(x_n)_{N=1}^{\infty}$ in $\mathbb{R}$ that is bounded above/*below* converges to some $x \in \mathbb{R}$

Definition: $\mathbb{R}^k_+ = \{(x_1, x_2, \ldots, x_k) \in \mathbb{R}^k \mid x_i \geq 0, \forall\, i \in \{1, 2, \ldots, k\}\}$

Definition: $\mathbb{R}^k_{++} = \{(x_1, x_2, \ldots, x_k) \in \mathbb{R}^k \mid x_i > 0, \forall\, i \in \{1, 2, \ldots, k\}\}$

Claim: Every non-empty set $S \subseteq \mathbb{R}$ that has an upper bound also has a least upper bound.

To fill the void of the rest of this page, here are some Willy Chen original puns:

Why does the heart only pump blood into the arteries?

Because it doesn't want its hard work to go in vein.

Why should you always bring a jacket to a brewery??

It gets drafty sometimes.

[7] This will end up being a super important fact that we see a lot in studying preferences

2 Metric Spaces

2.1 Metric Space Fundamentals

Consider the ordered field $(\mathbb{R}, +, \cdot, \geq)$. We need a little more structure to make this a metric space. Specifically, we need to define how distances are measured in this field. Conventionally, we use the following distance functions:

1. (In $\mathbb{R}^1$) $d(x, y) : \mathbb{R} \to \mathbb{R}$, $d(x, y) = |x - y|$
2. (Euclidean Distance in $\mathbb{R}^n$) $d(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$, $d(x, y) = \|x - y\| = \sqrt{\sum_i (x_i - y_i)^2}$

Definition (Metric): A **Metric** on a set X is a function $d : X \times X \to \mathbb{R}_+$ such that:

(i) **Separation** $d(x, y) = 0 \iff x = y$

(ii) **Symmetry** $d(x, y) = d(y, x)$

(iii) **Triangle Inequality** $\forall x, y, z \in X$, $d(x, y) \leq d(x, z) + d(z, y)$

Definition (Metric Space): A **Metric Space** M on the set X is a *pair* $M \equiv (X, d)$ where $d : X \times X \to \mathbb{R}_+$ is a metric

An example of a non-standard metric is **the discrete metric**. Consider the following distance function (metric):

$$d_0(x, y) = \begin{cases} 1 & , x \neq y \\ 0 & , x = y \end{cases}$$

One can check that the function $d_0(x, y)$ satisfies the 3 criteria for a distance function to be a metric. The following are the more "useful" metrics:

The p-metric: $\forall p \in [1, \infty)$ define the p-metric as the distance function d_p on $\mathbb{R}^n$ such that:

$$d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}}$$
$$d_\infty(x, y) = \sup_{i \in \{1, \ldots, n\}} \{|x_i - y_i|\}$$

Definition (sequence space l_p): The **Sequence Space** l_p is the space containing infinite real sequences that satisfy:

$$(x_k) \in l_p \Rightarrow \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} < \infty$$

$$(x_k) \in l_\infty \Rightarrow \sup_{i \in \{1,\dots,n\}} \{|x_i|\} < \infty$$[8]

These only discuss distances on sets with up to countably infinitely many points. For cases of distances with uncountably infinitely many points, we use the function spaces L^p. For example, let $\mathbf{C}[0,1]$ denote all continuous real functions on the interval [0,1]. $\forall\, p \in [1, \infty)$, $f, g \in \mathbf{C}[0,1]$, we have

$$d_p(f,g) \equiv \left(\int_0^1 |f(t) - g(t)|^p dt\right)^{\frac{1}{p}}$$

$$d_\infty(f,g) \equiv \underbrace{\max_{t\in[0,1]} |f(t) - g(t)|}_{\text{Since } f,g \text{ are continuous, this is equivalent to taking the supremum}}$$

Proposition: P-metric is valid $\forall p \in [1, \infty)$ on $\mathbb{R}^n$ and the l_p space.

Lemma 2.1: Holder's Inequality

Let (S, Σ, μ) be a measure space and let $p, q \in [1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then for all measurable real-or-complex-valued functions f and g on S

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

Theorem 2.1: Minkowski's Inequality

$\forall p \in [1, \infty)$, $\forall n \in \mathbb{N}$, $(x_k)_{k=1}^{\infty}, (y_k)_{k=1}^{\infty} \in \mathbb{R}^\infty$, $i = 1, \dots, n$,

$$\left(\sum_{i=1}^{n} |x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |x_k|^p\right)^{\frac{1}{p}}$$

Proof 2.1: Minkowski's Inequality (Efe Ok)

Notice that $\forall p \in [1,\infty)$, $f(x) = x^p$ is a convex function on $\mathbb{R}_+$ $(\frac{d^2 f}{dx^2} = p(p-1)x^{p-2} \geq 0)$, so $\forall\, a, b \in \mathbb{R}_+$, $f(\frac{a+b}{2}) \leq \frac{f(a)+f(b)}{2}$.

Take $(x_k), (y_k) \in \mathbb{R}^n$, we have that $\left(\sum\limits_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} < \infty$ and $\left(\sum\limits_{k=1}^n |y_k|^p\right)^{\frac{1}{p}} < \infty$, since we have finitely many dimensions.

Let $a = \left(\sum\limits_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}$, and $b = \left(\sum\limits_{k=1}^n |y_k|^p\right)^{\frac{1}{p}}$

Additionally, define $(\hat{x}_k) = \frac{1}{a}(x_k)$ and $(\hat{y}_k) = \frac{1}{b}(y_k)$, then we have for the i^{th} term:

$$|x_i - y_i|^p \leq (a|\hat{x}_i| + b|\hat{y}_i|)^p = (a+b)^p \left(\frac{a}{a+b}|\hat{x}_i| + \frac{b}{a+b}|\hat{y}_i|\right)^p \leq (a+b)^p \frac{(a|\hat{x}_i|^p + b|\hat{y}_i|^p)}{a+b}$$

$$\iff \sum_{i=1}^n (|x_i - y_i|^p) \leq (a+b)^{p-1} \cdot \left[a \cdot \underbrace{\sum_{i=1}^n (|\hat{x}_i|^p)}_{=1} + b \cdot \underbrace{\sum_{i=1}^n (|\hat{y}_i|^p)}_{=1}\right] = (a+b)^{p-1} \cdot (a+b)$$

$$\iff \left[\sum_{i=1}^n (|x_i - y_i|^p)\right]^{\frac{1}{p}} \leq a + b = \left[\sum_{i=1}^n |x_i|^p\right]^{\frac{1}{p}} + \left[\sum_{i=1}^n |y_i|^p\right]^{\frac{1}{p}}$$

Remark: This proof is an excerpt from Efe Ok p.125-126. Alternatively, you may see on google that you can prove it with Holder's Inequality. $\square$

Definition (Distance): Given a metric space $M \equiv (X, d_x)$ and $x \in X, A \subseteq X$, we define the **Distance** between a point $x \in X$ and a set $A \subseteq X$ to be:

$$d_X(x, A) \equiv \inf_{a \in A} d_X(x, a)$$

Definition: Given a metric space $M \equiv (X, d_x)$, $\forall x_0 \in X, \forall\, \varepsilon \in \mathbb{R}_{++}$, the $\varepsilon-$neighborhood of x_0 is defined as:

$$N_\varepsilon^{d_X}(x_0) \equiv \{x \in X \mid d(x, x_0) < \varepsilon\}$$

Definition: Given a metric space $M \equiv (X, d_x)$, $\forall A \subset X, \forall \varepsilon \in \mathbb{R}_{++}$, the $\varepsilon-$neighborhood of A is defined as:

$$N_\varepsilon^{d_X}(A) \equiv \{x \in X \mid \exists a \in A,\, d(x, x_0) < \varepsilon\} = \bigcup_{a \in A} N_\varepsilon^{d_X}(a)$$

Definition (Open Set): A metric space $M \equiv (X, d_x)$, $\forall O \subseteq X$, we say that O is **Open** in X with respect to d_X if $\forall x \in O$, $\exists \varepsilon \in \mathbb{R}_{++}$, $N_{\varepsilon}^{d_X}(x) \subseteq O$

Definition (Closed Set): A metric space $M \equiv (X, d_x)$, $S \subseteq X$ is **Closed** if $X \setminus S$ is open.

Equivalent definition for closed-ness: Given a metric space $M \equiv (X, d_x)$, $S \subseteq X$ is closed if every convergent sequence in S converges to a point in S.

Set Operations and Open/Closed-ness:

- Every *union* of *open sets* is open
- Every *intersection* of *closed sets* is closed
- Every finite *union* of *closed sets* is closed
- Every finite *intersection* of *open sets* is open

Proof: Set Operations and Open/Closed-ness

Every union of open sets is open.
Suppose otherwise that this is not true and that there exists some open sets $A, B \subset X$ such that $A \cup B$ is not open. Then $\exists a \in A \cup B$ such that $\forall \varepsilon > 0$, $\exists c \in A^c \cap B^c$ such that $c \in N_{\varepsilon}^{d}(a)$. Since $a \in A \cup B$, a is in at least one of A or B. and c is in neither A or B. But that means whichever set a is from, that set is not an open set, which is a contradiction.

Every intersection of closed sets is closed.
Given two closed set $A, B \subset X$, their complements must be open. i.e., A^c, B^c are open. Since A^c, B^c are open, their union is also open. i.e.,$A^c \cup B^c = (A \cap C)^c$ is open. Since $(A \cap C)^c$ is open, by definition, $A \cap C$ is closed.

Every finite intersection of open sets is open
Take n open sets $A_1, A_2, \ldots, A_n \subseteq X$, we first show that $\bigcap\limits_{i=1}^{n} A_i$ is open. i.e., the finite intersection of open sets is open.
Since all A_i, $i \in \{1, 2, \ldots, n\}$ are open, we know that $\forall a_i \in A_i$, $\exists\, \varepsilon_i > 0\, s.t.\, N_{\varepsilon_i}^{d}(a_i) \subseteq A_i$. Then $\forall a \in \bigcap\limits_{i=1}^{n} A_i$ we know $a \in A_i, \forall i \in \{1, 2, \ldots, n\}$. Take $\varepsilon^* = \min\limits_{i \in \{1,2,\ldots,n\}} \varepsilon_i$, then we

know that $N^d_{\varepsilon^*}(a) \subseteq \bigcap\limits_{i=1}^{n} A_i$.

To show that the infinite case fails, take $A_i = (-\infty, \frac{1}{n}) \cup (1, \infty) \subset \mathbb{R}$, $n \in \mathbb{N}$, then $\bigcap\limits_{i=1}^{\infty} A_i = (\bigcup\limits_{i=1}^{\infty} A_i)^c \to \mathbb{R} \setminus (0,1] = (-\infty, 0] \cup (1, \infty)$. At $0 \in \bigcap\limits_{i=1}^{\infty} A_i$, take any $0 < \varepsilon < 1$, we can see that $0 + \varepsilon = \varepsilon \notin \bigcap\limits_{i=1}^{\infty} A_i$, hence the $\bigcap\limits_{i=1}^{\infty} A_i$ is not closed.

Every finite union of closed sets is closed

By definition of closed sets, their complement is open. Take n closed sets $A_1, A_2, \ldots, A_n \subseteq X$, we know that $A_1^c, A_2^c, \ldots, A_n^c \subseteq X$ are open, and we know, from the previous proof, that $\bigcap\limits_{i=1}^{n} A_i^c$ are open as well. Let the finite union of these closed sets be $A = \bigcup\limits_{i=1}^{n} A_i$, and its complement, by De Morgan's Law is $A^c = \bigcap\limits_{i=1}^{n} A_i^c$. Since A^c is open, by definition, the set $A = \bigcup\limits_{i=1}^{n} A_i$, the finite union of n closed sets, is closed.

To show that the infinite case fails, take $A_i = [\frac{1}{i}, 1]$, then $\bigcup\limits_{i=1}^{\infty} A_i \to (0,1]$. But the sequence $\frac{1}{n}$ converges to $0 \in (0,1]$, making the union a non-closed set.

$\square$

Definition: Given a metric space $M \equiv (X, d_x)$, $\forall Y \subseteq X$,

- $x \in X$ is a **boundary point** of Y if $\forall \varepsilon \in \mathbb{R}_{++}$, $N_\varepsilon^{d_X}(x) \cap Y \neq \varnothing \wedge N_\varepsilon^{d_x}(x) \cap Y^c \neq \varnothing$
- $x \in X$ is an **interior point** of Y if $\exists \varepsilon \in \mathbb{R}_{++}$, $N_\varepsilon^{d_X}(x) \subseteq Y$

Definition: Given a metric space $M \equiv (X, d_x)$. The **Interior** of $Y \subseteq X$ (denoted as $int(Y)$) is defined as (the union of all interior points):

$$int(Y) \equiv \bigcup \{O \subseteq Y \mid O \text{ is open}\}$$

Definition: Given a metric space $M \equiv (X, d_x)$. The **Closure** of $Y \subset X$ (denoted as $cl(Y)$) is defined as (the smallest closed set that contains Y):

$$cl(Y) \equiv \bigcap \{S \subseteq X \mid Y \subseteq S, S \text{ is closed}\}$$

Alternatively, $cl(Y) = Y \cup \{\text{limit points of } Y\}$.

Definition (Boundary): Given a metric space $M \equiv (X, d_x)$. The **Boundary** of $Y \subseteq X$ (denoted as $bd(Y)$) is defined as:

$$bd(Y) = cl(Y) \setminus int(Y)$$

Sequential Definition of a Limit: Given a metric space $M \equiv (X, d_x)$, $(x_k)_{k=1}^{\infty}$ on X, (x_k) is said to converge to x ($\lim\limits_{k\to\infty} x_k = x \in X$) if $\forall \varepsilon \in \mathbb{R}_{++}, \exists m \in \mathbb{N}$ $s.t.$ $\forall k \geq m$, $d(x_k, x) < \varepsilon$

Definition: Given a metric space $M \equiv (X, d_x)$ the lim sup and lim inf of a sequence $(x_k)_{k=1}^{\infty}$ are defined as:

$$\limsup (x_k)_{k=1}^{\infty} \equiv \lim_{n\to\infty} (\sup\{x_k \mid k \geq n\})$$
$$\liminf (x_k)_{k=1}^{\infty} \equiv \lim_{n\to\infty} (\inf\{x_k \mid k \geq n\})$$

Theorem 2.2: Sequence Convergence

Given a metric space $M \equiv (X, d_x)$. $(x_k)_{k=1}^{\infty}$ is convergent if and only if

$$\limsup(x_k) = \liminf(x_k) = \lim_{k\to\infty}(x_k)$$

Theorem 2.3: Limit Point is Unique

Given a metric space $M \equiv (X, d_x)$.

$$\lim_{k\to\infty}(x_k) = x \in X \wedge \lim_{k\to\infty}(x_k) = y \in X \iff x = y$$

Proof 2.3

Suppose otherwise that there exists a convergence sequence $\{x_n\} \in \mathbb{R}^n$ that has 2 limit points x and x'. Then $\forall \varepsilon > 0, \exists N, N' \in \mathbb{N}$ such that $\forall n > N, \|x_n - x\| < \varepsilon$ and $\forall n > N', \|x_n - x'\| < \varepsilon$.
Take $\varepsilon^* = \frac{1}{4}\|x - x'\|$ and $N^* = \max\{N, N'\}$. Then we have $\forall n \geq N^*$, $\|x_n - x\| < \varepsilon^* = \frac{1}{4}\|x - x'\| = \frac{1}{4}\|x - x_n + x_n - x'\| \leq \frac{1}{4}\|x_n - x\| + \frac{1}{4}\|x_n - x'\| < \frac{1}{2}\varepsilon^*$
Since $\varepsilon^* > 0$, $\varepsilon^* < \frac{1}{2}\varepsilon^*$ is a contradiction. Hence a convergent sequence in $\mathbb{R}^n$ can only have one limit point. $\square$

Theorem 2.4: Alternative Definition of Closedness

Given a metric space $M \equiv (X, d_x)$, $Y \subseteq X$ is closed if and only if every sequence in Y that converges in X also converges in Y. i.e.,

$$Y \text{ is closed } \iff [(y_k)_{k=1}^{\infty} \in Y^{\infty}, ((y_k) \to x \in X) \;\Rightarrow\; (x \in Y)]$$

Proof 2.4

" $\Rightarrow$ "

Assume that Y is closed. Take the sequence $(x_k)_{k=1}^{\infty} \in Y^{\infty}$ with $(x_k) \to x \in X$. Suppose otherwise that $x \in X \setminus Y$, then $X \setminus Y$ is closed and $\exists \varepsilon \in \mathbb{R}_{++}$ $s.t.$ $N_{\varepsilon}^{dx}(x) \subseteq X \setminus Y$. Since $(x_k) \to x$, $\lim_{k \to \infty} d(x_k, x) = 0$, $\exists n \in \mathbb{N}$, $x_n \in N_{\varepsilon}^{dx}(x)$, so $x_n \notin Y$, but by definition, $x_n \in Y$, so by contradiction, $x \in Y$.

"$\Leftarrow$"

Assume that for every sequence $(x_k)_{k=1}^{\infty} \in Y$, we have $(x_k)_{k=1}^{\infty} \to x \in X \;\Rightarrow\; x \in Y$. Suppose otherwise that Y is not closed, then $X \setminus Y$ is not open. Take $x \in X \setminus Y$ such that $\forall \varepsilon \in \mathbb{R}_{++}, N_{\varepsilon}^{dx}(x) \cap Y \neq \varnothing$. Then $\forall m \in \mathbb{N}$, $\exists x_m \in N_{\varepsilon m}^{dx}(x) \cap Y$. So $(x_m)_{m=1}^{\infty} \in Y^{\infty}$, and by assumption, $(x_m)_{m=1}^{\infty} \to x \in Y$. But by construction, $x \in X \setminus Y \iff x \notin Y$. Hence, by contradiction, Y has to be closed.[a] $\square$

[a]The gist of this proof is that "If Y is not closed, then no limit $x \in Y$. Since we found that $x \in Y$, then Y must be closed.

Important: Open/Closed-ness always depends on the underlying metric space (X, d_X). For example, $(0, 1)$ is open in $(\mathbb{R}, d_1)$ but closed in $((0, 1), d_1)$

Important: Open/Closed-ness is not binary. Sets can be neither open nor closed; sets can also be both open and closed.

2.2 Properties of Well-Behaved Metric Spaces

2.2.1 Connectedness

Definition: Given a metric space $M \equiv (X, d_x)$. M is said to be **Connected** if a subspace *cannot* be obtained without cutting the space [9]

[9]For example, $[0, 1]$ is connected, but $[0, 1] \cup [2, 3]$ is not connected

Definition: Given a metric space $M \equiv (X, d_x)$. M is **Connected** if there does *not* exist 2 non-empty, disjoint, and open subsets A, B such that $A \cup B = X$

Definition: Given a *connected* metric space $M \equiv (X, d_x)$, a subset $Y \subseteq X$ is connected in X if Y is a connected metric subspace of X

2.2.2 Separability

Definition (Dense): Given a metric space $M \equiv (X, d_x)$. $Y \subseteq X$ is **Dense** in X if $cl(Y) = X$

Definition (Separable): Given a metric space $M \equiv (X, d_x)$. X is **Separable** if X contains a subset that is countable and dense.

Theorem 2.5: Weierstrass Approximation Theorem

$\forall\, a, b \in \mathbb{R}$, the set of all polynomial functions on $[a, b]$ is dense in $\mathbf{C}[0, 1]$

Corollary: $\mathbf{C}[a, b]$, the set of all continuous functions on $[a, b]$, is separable

Proof 2.5

The set of rational polynomials is countable since there are finitely many terms (for a polynomial) and the coefficients are rational. The closure will include irrational coefficient polynomials given completeness in $\mathbb{R}$, so $\mathbf{C}[0, 1]$ is dense in $\mathbb{R}$

$\square$

Theorem 2.6

Given a metric space $M \equiv (X, d_x)$ is separable. There exists a countable class $\mathbb{O}$ of open sets in X such that $\forall$open $U \subseteq X, U = \bigcup\{O \in \mathbb{O} \mid O \subset U\}$

Proof 2.6

Since (X, d_X) is separable, take $Y \subseteq X$ be a countable and dense subset of X. Define $\mathbb{O} \equiv \{N_\varepsilon(z) \mid z \in Y, \varepsilon \in \mathbb{Q}_{++}\}$ (Notice that $\mathbb{O}$ is a countable set of open sets).

Take an open subset $U \subseteq X$ and $x \in U$, we want to show that $x \in O$ for some $O \in \mathbb{O}$ such that $O \subseteq U$. Since U is open, $\exists \varepsilon \in \mathbb{Q}_{++}$ such that $N_\varepsilon(x) \subseteq U$. Then since $cl(Y) = X$ (because (X, d_x) is separable), $\exists y \in Y$ with $d(x,y) < \frac{\varepsilon}{2}$. i.e., $x \in N^{d_X}_{\frac{\varepsilon}{2}}(y) \subseteq N^{d_X}_{\varepsilon}(x) \subseteq U$. Since $y \in Y \wedge x \in N^{d_X}_{\frac{\varepsilon}{2}}(y)$, we know $x \in N^{d_X}_{\frac{\varepsilon}{2}}(y) = O \in \mathbb{O}$

$\square$

2.2.3 Completeness

Definition: Given a metric space $M \equiv (X, d_x)$. The sequence $(x_k)_{k=1}^{\infty}$ is a **Cauchy Sequence** if $\forall \varepsilon \in \mathbb{R}_{++}$, $\exists m_\varepsilon \in \mathbb{N}$ such that $\forall\, j,k \in \{h \in \mathbb{N} \mid h > m_\varepsilon\}$, $d_X(x_j, x_k) < \varepsilon$

Definition (Completeness): Given a metric space $M \equiv (X, d_x)$. M is **Complete** if every Cauchy sequence in X converges in X

2.2.4 Boundedness

Definition3 (Bounded): Given a metric space $M \equiv (X, d_x)$. The set $S \subseteq X$ is said to be **Bounded** if $\exists \varepsilon \in \mathbb{R}_{++}, x \in X \ s.t.\ X \subseteq N^{d_X}_{\varepsilon}(x)$

2.2.5 Compactness

Definition: Given a metric space $M \equiv (X, d_x)$, $Y \subseteq X$. A collection $\mathbb{O}$ of (open) subsets of X is said to be a(n) **Open Cover** of Y if $Y \subseteq \bigcup\limits_{o_i \in \mathbb{O}} o_i$

Definition (Compactness): Given a metric space $M \equiv (X, d_x)$. M is compact if *every open cover* of X has a finite subset that is also an open cover of X

Definition: Given a metric space $M \equiv (X, d_x)$, $Y \subseteq X$ is compact in X if every open cover of Y has a finite open sub-cover[10]

[10] Note that this is NOT equivalent to having a finite open cover, since the X is itself a finite open cover.

Definition: Given a metric space $M \equiv (X, d_x)$, $Y \subset X$ is **Sequentially Compact** if every sequence in Y has a subsequence that converges to a point in Y

Theorem 2.7

$Y \subseteq X$ is compact if and only if Y is sequentially compact in X

Theorem 2.8

$Y \subseteq X$ is compact $\Rightarrow$ Y is closed and bounded.

Proof 2.8

Assuming that Y is compact, we need to show that $X \setminus Y$ is open (so Y is closed). If $X \setminus Y = \varnothing$, then closed and bounded is trivially true or false since we don't know anything about X. So let's assume that $Y \subset X$ so that $X \setminus Y \neq \varnothing$
Take $x \in X \setminus Y$, then $\forall y \in Y$, $\exists\, \varepsilon_y = \frac{d_X(x,y)}{2} \in \mathbb{R}_{++}$ such that $N^{d_X}_{\varepsilon_y}(x) \cap N^{d_X}_{\varepsilon_y}(y) = \varnothing$. Moreover, since $\{N^{d_X}_{\varepsilon_y}(y) \mid y \in Y\}$ is an open cover of Y and Y is compact, we know that there is a finite open subcover Z of $\{N^{d_X}_{\varepsilon_y}(y) \mid y \in Y\}$ such that $\{N^{d_X}_{\varepsilon_y}(y) \mid y \in Z, |Z| < \infty\}$ also covers Y. Now define $\varepsilon^* = \min_{y \in Z} \varepsilon_y$, then $N^{d_X}_{\varepsilon}(x) \subseteq X \setminus Y$, hence $X \setminus Y$ is open and so Y is closed. Since Y is compact and not equal to X, Y must also be bounded because $Y \subseteq N^{d_X}_{\tilde{\varepsilon}}(y), \varepsilon = \sum \varepsilon^*, y \in Y$ □

Theorem 2.9: A Closed Subset of a Compact Space is Compact

Let $M \equiv (X, d_x)$ be a compact metric space. If $Y \subseteq X$ is closed, Y is compact.

Proof 2.9: Heine-Borel Theorem

Take $\mathbb{O}$ to be an open cover of Y, then $\mathbb{O} \cup \{X \setminus Y\}$ is an open cover of X [a]. Since X is compact, we know that $\mathbb{O} \cup \{X \setminus Y\}$ has a finite open subcover $\mathbb{O}'$ of X. Since $\mathbb{O}'$ is finite, $\mathbb{O}' \setminus \{X \setminus Y\}$ is also finite. By construciton, $\mathbb{O}' \setminus \{X \setminus Y\}$ is an open subcover of $\mathbb{O}$ of Y. So $\mathbb{O} \cup \{X \setminus Y\}$ is a finite open subcover of Y and hence Y is compact. □

[a] Since Y is closed, $X \setminus Y$ is open.

Theorem 2.10

In the standard $(\mathbb{R}^n, d_2)$ metric space, take any set $Y \subset \mathbb{R}^n$, then

$$Y \text{ is compact} \iff Y \text{ is closed and bounded}$$

Proof 2.10: Heine-Borel Theorem

First we need to show that if $Y = \mathbb{R}^n$, Y cannot be compact in $(\mathbb{R}^n, d_2)$. Take the sequence of open sets $(-k,k)^n \subset \mathbb{R}^n$, then $\bigcup\limits_{k=1}^{\infty}(-k,k)^n$ is open, and more importantly, an open cover of $\mathbb{R}^n$. Notice that $\bigcup\limits_{k=1}^{\infty}(-k,k)^n$ does not have any finite collection of subsets that also cover $\mathbb{R}^n$ as $\bigcup\limits_{k=1}^{\alpha}(-k,k)^n$, $\alpha < \infty$ does not cover $\mathbb{R}^n$. Hence we only need to prove the statement for $Y \subset \mathbb{R}^n$

"$\Rightarrow$":
Assume that $S \subset \mathbb{R}^n$ is compact in $(\mathbb{R}^n, d_2)$. We want to show that $\mathbb{R}^n \setminus S$ is an open set and that S is bounded. Take $x \in \mathbb{R}^n \setminus S$ and $s \in S$, and take $\varepsilon_s = \frac{d(x,y)}{2}$, then we know $N^{d_2}_{\varepsilon_s}(x) \cap N^{d_2}_{\varepsilon_S}(s) = \varnothing$. We also know that $\{N^{d_2}_{\varepsilon_s}(s) \mid s \in S\}$ is an open cover for S. Since S is compact, $\exists \mathbb{O} \subset \{N^{d_2}_{\varepsilon_s}(s) \mid s \in S\}$, $|\mathbb{O}| < \infty$, $S \subseteq \bigcup \mathbb{O}$. Take $\varepsilon^* = \min\limits_{o \in \mathbb{O}} \varepsilon_o$, then $N^{d_2}_{\varepsilon_o}(x) \subset N^{d_2}_{\varepsilon_s}(x) \subseteq \mathbb{R}^n \setminus S$. Since $x \in \mathbb{R}^n \setminus S$ is taken arbitrarily, we know $\mathbb{R}^n \setminus S$ is open and thus S is **closed**. Moreover, take $\varepsilon' = \sum\limits_{o \in \mathbb{O}}$, since $\mathbb{O}$ is finite, we know $\varepsilon' < \infty$ and that $\forall\, s \in S$, $S \subseteq N^{d_2}_{e'}(s)$, so S is also **bounded**.

"$\Leftarrow$":
Assume that S is closed and bounded. Now supposed otherwise that S is not compact, then it is not sequentially compact. i.e., $\exists (s_n) \in S$ such that all subsequences $(s_{n_i}) \in S$ either does not converge or converge outside of S. But that would mean there is a sequence in S that converges outside of S, so S is not a closed set. Moreover, suppose that S is closed but not compact. Since S is not compact, there exists an open cover $\mathbb{O}'$ for S that does not have a finite set of open sub-covers. So $\exists s' \in S, \forall k < \infty, s' \notin \bigcup_{i=1}^{k}\{o_i \in \mathbb{O}'\}$. In other words, $\forall \varepsilon > 0, S \nsubseteq N^{d_2}_{\varepsilon}(s')$, meaning that S is not bounded. $\square$

2.2.6 Continuity of Functions

Definition: Given two metric spaces (X, d_x), (Y, d_Y), and the function $f : X \to Y$. We say that f is **Continuous** on X at $x_0 \in X$ if $\forall \varepsilon \in \mathbb{R}_{++}$, $\exists \delta_\varepsilon(x_0) \in \mathbb{R}_{++}$ such that

$$\forall x \in X, d_X(x, x_0) < d_\varepsilon(x_0) \ \Rightarrow \ d_Y(f(x), f(x_0)) < \varepsilon$$

or equivalently,

$$\forall x \in N_{\delta_\varepsilon}^{d_X}(x_0),\ f(x) \in N_\varepsilon^{d_Y}(f(x_0))$$

Equivalent Definition: Given a metric space $(X, d_x), (Y, d_Y)$. A function $f : X \to Y$ is continuous at $x_0 \in X$ if and only if $\forall$ open $O \subseteq Y\, s.t.\, f(x_0) \in O$, $\exists \delta \in \mathbb{R}_{++}\, s.t.\, \forall x \in N_\delta^{d_X}(x_0),\ f(x) \in O$

Definition: Given a metric space $M \equiv (X, d_x)$. $f : X \to \mathbb{R}$ is **Upper-Semi-Continuous** at $x_0 \in X$ if $\forall \varepsilon \in \mathbb{R}_{++}, \exists \delta_\varepsilon \in \mathbb{R}_{++}$ such that

$$\forall x \in X,\ d_X(x, x_0) < \delta_\varepsilon \ \Rightarrow \ f(x) \leq f(x_0) + \varepsilon$$

or equivalently,

$$\limsup_{x \to x_0} f(x) \leq f(x_0)$$

Equivalent Definition: Given a metric space $M \equiv (X, d_x)$. $f : X \to \mathbb{R}$ is **Upper-Semi-Continuous** at $x_0 \in X$ if the set $\{x \in X \mid f(x) \geq x_0\}$ is closed.

Definition: Given two metric space $(X, d_X), (Y, d_Y)$. $f : X \to \mathbb{R}$ is **Uniformly Continuous** if $\forall \varepsilon \in \mathbb{R}_{++}, \exists \delta_\varepsilon \in \mathbb{R}_{++}$ such that:

$$\forall x, x_0 \in X, d_X(x, x_0) < \delta_\varepsilon \ \Rightarrow \ d_Y(f(x), f(x_0)) < \varepsilon$$

2.2.7 Continuity of Correspondences

Recall the open set definition of functional continuity. We want to keep using a similar definition but expand to correspondences. What can we do?

Definition (Upper Hemi-Continuity): Given metric spaces $(X, d_x), (Y, d_Y)$. The correspondence $g : X \rightrightarrows Y$ is **Upper-Hemi-Continuous (*uhc*)** at $x_0 \in X$ if and only if $\forall$ open $O \subseteq Y\ s.t.\ g(x) \subseteq O, \exists\, \delta_O \in \mathbb{R}_{++}\, s.t.\, g(N_{\delta_O}^{d_X}(x_0)) \subseteq O$

Definition: Given metric spaces $(X, d_x), (Y, d_Y)$. The correspondence $g : X \rightrightarrows Y$ is **Closed-Valued** at $x_0 \in X$ if $\forall x \in X,\ g(x) \subseteq Y$ is a closed set.

Definition: Given metric spaces $(X, d_x), (Y, d_Y)$. The correspondence $g : X \rightrightarrows Y$ is **Compact-Valued** at $x_0 \in X$ if $\forall x \in X,\ g(x) \subseteq Y$ is a compact set.

Proposition: Given metric spaces $(X, d_x), (Y, d_Y)$, the *compact-valued* correspondence $g : X \rightrightarrows Y$, and $x_0 \in X$. If for every sequence $(x_k)_{k=1}^{\infty} \to x_0$ and every sequence $(y_k)_{k=1}^{\infty}$ such that $y_k \in g(x_k)$, there exists a convergent subsequence $(y_{k_j})_{j=1}^{\infty} \to y \in g(x_0)$, then g is *upper-hemi-continuous* at $x_0 \in X$ (See Efe Ok p.288 for pictoral representation)

Definition: Given metric spaces $(X, d_x), (Y, d_Y)$ and the correspondences $g : X \rightrightarrows Y$. The **Graph** of g is the set $gr(g) \equiv \{(x, y) \in X \times Y \mid y \in g(x)\} \subseteq X \times Y$

Definition: Given metric spaces $(X, d_x), (Y, d_Y)$ and the correspondences $g : X \rightrightarrows Y$. We say that g has a **Closed Graph** in $X \times Y$ if $gr(g)$ is closed in $(X \times Y, d_{X\times y})$

Remark: Given metric spaces $(X, d_x), (Y, d_Y)$. The standard metric in Cartesian product is called the **Product-Metric** and is defined as

$$\begin{aligned} &d_{X\times Y} : (X \times Y) \times (X \times Y) \to \mathbb{R}_+, \\ &d_{X\times Y}\left((x, y), (x', y')\right) = d_X(x, x') + d_Y(y, y') \end{aligned}$$

This is analogous to using d_1 on $\mathbb{R}^2$.

Definition: Given metric spaces $(X, d_x), (Y, d_Y)$ and the correspondences $g : X \rightrightarrows Y$. We say that $g : X \rightrightarrows Y$ is **Closed** at $x_0 \in X$ if $\forall \left((x_k, y_k)\right)_{k=1}^{\infty}, x_k \in X, y_k \in g(x_k), \forall k \in \mathbb{N}, \left((x_k, y_k)\right)_{k=1}^{\infty} \to (x_0, y_0)$, we have $y_0 \in g(x_0)$. In other words, a closed graph is a graph of g that is closed at every point in the graph.[11]

Proposition: Given metric spaces $(X, d_x), (Y, d_Y)$ and the correspondence $g : X \rightrightarrows Y$. If $g : X \rightrightarrows Y$ is **uhc AND closed-valued**, then it has a closed graph.

Proposition: Given metric spaces $(X, d_x), (Y, d_Y)$ and the correspondence $g : X \rightrightarrows Y$. If Y

[11] Being closed and being closed-valued are not equivalent.

is compact and g has a closed graph, then g is **upper-hemi-continuous** everywhere on X.
Definition (Lower Hemi-Continuity: Given metric spaces $(X, d_x), (Y, d_Y)$. The correspondences $g : X \rightrightarrows Y$ is **Lower-Hemi-Continuous** at $x_0 \in X$ if

$$\forall y \in g(x_0), \forall (x_k)_{k=1}^{\infty} \to x_0 \text{ with } x_k \in X, \forall k \in \mathbb{N},\ \exists (y_k)_{k=1}^{\infty} \to y_0\ s.t.\ y_k \in g(x_k),\ \forall k \in \mathbb{N}$$

In other words, g is lower-hemi-continuous at $x_0 \in X$ if for all $y \in g(x_0)$, every sequence in X that converges to x_0 has a corresponding sequence in Y that converges to y. Notice that this is very different from the sequential characterization of *uhc* because it requires **every point** in the image to have convergent sequences from **all directions** of the domain.

Equivalent Definition (Open sets): Given metric spaces $(X, d_x), (Y, d_Y)$. The correspondence $g : X \rightrightarrows Y$ is **Lower-Hemi-Continuous** at $x_0 \in X$ if

$$\forall \text{ open } O \subseteq Y \text{ such that } g(x_0) \cap O \neq \varnothing \ \Rightarrow\ \forall x \in N_{\varepsilon}^{d_X}(x_0),\ g(x) \cap O \neq \varnothing$$

Definition: Given metric spaces $(X, d_x), (Y, d_Y)$. The correspondences $g : X \rightrightarrows Y$ is **Continuous** at $x_0 \in X$ if and only if it is *both* uhc and lhc at x_0.

2.3 Slight Detour for Economic Motivations

Our typical economics problem is to maximize a function subject to constraints. In these settings, having a continuous function f over a compact set X makes things very convenient.

Proposition: Given metric spaces $(X, d_x), (Y, d_Y)$ and a continuous function $f : X \to Y$. If $Z \subseteq X$ is a compact set, then $f(Z) \subseteq Y$ is also a compact set.

Theorem 2.11: Weierstrass Maximum Theorem

Given a metric space $M \equiv (X, d_x)$ and $A \subseteq X$ a compact subset of X. If $f : A \to \mathbb{R}$ is a continuous function, then $\exists a_{max}, a_{min} \in A$ such that $f(a_{max}) = \sup\{f(a)\}$ and $f(a_{min}) = \inf\{f(a)\}$. Notationally, a_{min} and a_{max} are defined as:

$$a_{min} \equiv \underset{a \in A}{\text{argmin}} f(a)$$
$$a_{max} \equiv \underset{a \in A}{\text{argmax}} f(a)$$

In other words, the *Weierstrass Maximum Theorem* says that a function is continuous over a compact set, there exists maximizers and minimizers of the function that obtains extreme values within the compact range of the function.

This theorem allows to know that if we have a compact domain (think budget set) and a continuous utility function, then we must have a utility maximizing bundle.

Proof 2.11: Weierstrass Maximum Theorem

To show that $\exists a_{max} \in A$ such that $f(a_{max}) = \sup\{f(a)\}$, construct a sequence $(a_k)_{k=1}^{\infty}$ such that $\forall i, j \in \mathbb{N}$ $s.t.$ $a_i \neq a_j, i < j, \iff f(a_i) \leq f(a_j)$ (i.e., construct a sequence so that the sequence $f(x_k)$ is strictly increasing). Then since A is assumed to be compact, there is a convergent subsequence $(a_{ki})_{i=1}^{\infty}$ of $(a_k)_{k=1}^{\infty}$ that converges to a point $a \in A$. Since f is continuous, we know that $f(a_{ki}) \to f(a) \in f(A) = \{x \in \mathbb{R} \mid x = f(a), a \in A\}$. By the construction of the sequence $(a_k)_{k=1}^{\infty}$, we know that $\sup(\{f(a)\} = f(a) \in f(A)$ and $a = a_{max}$

Similarly, to show that $\exists a_{min} \in A$ such that $f(a_{min}) = \inf(\{f(a)\}$, we construct a sequence $(a'_k)_{k=1}^{\infty}$ such that $\forall i, j \in \mathbb{N}$ $s.t.$ $a_i \neq a_j, i < j \iff f(a'_i) \geq f(a'_j)$. Then since A is assumed to be compact, there is a convergent subsequence $(a'_{ki})_{i=1}^{\infty}$ of $(a'_k)_{k=1}^{\infty} \to a' \in A$. Since f is continuous, we know that $f(a'_{ki}) \to f(a') \in f(A) = \{x \in \mathbb{R} \mid x = f(a), a \in A\}$. By the construction of the sequence $(a'_k)_{k=1}^{\infty}$, we know that $sup(\{f(a)\} = f(a') \in f(A)$ and $a' = a_{min}$ $\square$

Theorem 2.12: Berge's Theorem of Maximum (Proof)

Let $X, Y \subseteq \mathbb{R}^n$ be non-empty sets. Let Y be a compact set and $f : X \times Y \to R$ be a continuous function. Let $V(x) \equiv \max_{y \in Y} f(x, y)$ and $y^*(x) = \operatorname{argmax}_{y \in Y} f(x, y)$, then
$V : X \to \mathbb{R}$ is continuous and $y^*(x)$ is upper-hemi-continuous.

Why did the police release printed pictures of the suspects but will never arrest them.

They were framed.

3 Linear/Vector Spaces

Notice that in our textbook (De la Fuente), the convention is to use vector group $(V, +)$ over **any** field $(F, +, \cdot)$ to characterize a linear/vector space. However, for our purposes, it suffices to just use the vector group $(V, +)$ and the real field $(\mathbb{R}, +, \cdot)$. We will soon define a vector space to be $((V, +), \cdot_{\mathbb{R}})$.

3.1 Foundation of Linear Spaces

Definition (Field): A **Field** F is an abelian/commutative group $(F, *)$ endowed with a second operation $\circ$ that has $\circ-$inverse elements for all elements of F other than the $*$-identity and is distributive so that $x \circ (y * z) = x \circ y * x \circ z$

Definition: $(V, +, \cdot)$, $\underbrace{+ : V \times V \to V}_{\text{Vector Addition}}$, $\underbrace{\cdot : \mathbb{R} \times V \to V}_{\text{Scalar Multiplication}}$ is a **Vector/Linear Space** if $\forall \lambda, \lambda' \in \mathbb{R},\ u, v \in V$,:

(i) $(V, +)$ is an abelian group and $(V, \cdot)$ is closed

(ii) (Double Distributive Property Among $(+, \cdot)$): $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$ and $(\lambda + \lambda') \cdot u = \lambda u + \lambda' \cdot v$

(iii) $\cdot$ is associate: $\lambda(\lambda' \cdot v) = \lambda\lambda' \cdot v = \lambda' \cdot (\lambda \cdot v)$

(iv) (Neutrality of Scalar Identity 1): $1 \cdot v = v$

Properties of a linear space $(V, +, \cdot)$ **with the field** $(\mathbb{R}, +, \cdot)$

(i) $\vec{0} \in V$ is unique

(ii) $\forall \lambda \in \mathbb{R},\ \lambda \cdot \vec{0} = \vec{0}$

(iii) $\forall v \in V,\ 0 \cdot v = \vec{0}$

(iv) $\lambda v = \vec{0}$ if and only if $\lambda = 0$ or $v = \vec{0}$

(v) The additive inverse $-v$ is unique for all v in V

(vi) $\forall v \in V,\ (-1) \cdot v = -v$

(vii) $\forall \lambda \in \mathbb{R}, v \in V,\ (-\lambda)v = \lambda(-v) = -(\lambda v)$

Theorem 3.1: The set of all real functions is a linear space

Let X be a non-empty set and F be the set of all real functions on X. Then $(F, +, \cdot)$ with vector addition $(+)$ and scalar multiplication $(\cdot)$ such that $\forall f, g \in F, x \in X, \lambda \in \mathbb{R}$:

$$(f+g)(x) = f(x) + g(x)$$
$$(\lambda \cdot f)(x) = \lambda \cdot f(x)$$

F is a linear space.

Definition: Let $(V, +, \cdot)$ be a linear space. A subset $S \subseteq V$ is a **Linear Subspace** of V if $(S, +, \cdot)$ is a linear space.

Proposition: Let $(V, +, \cdot)$ be a linear space. A non-empty subset $S \subseteq V$ is a linear subspace of V if and only if $\forall \lambda \in \mathbb{R}, x, y \in S$, $\lambda x + y \in S$.

Proof:

" $\Rightarrow$ "
Since S is a linear space, $\lambda x \in S$ and $x + y \in S$, hence $\lambda x + y \in S$.

"$\Leftarrow$"
Assume that $\forall x, y \in S$, $\lambda \in \mathbb{R}$, $\lambda x + y \in S$. We need to show that:

(1) S is closed under vector addition and scalar multiplication.
Let $\lambda = 1$, then $\forall x, y \in S$, $1 \cdot x + y = x + y \in S$
Let $y = \vec{0}$, then $\forall x, y \in S$, $\lambda x + \vec{0} = \lambda x \in S$[a].

(2) S contains the additive identity.
Let $x = y$ and $\lambda = (-1)$, then $\lambda x + y = (-1)x + x = -x + x = \vec{0} \in S$

(3) S contains the additive inverses of elements in S.
Let $y = \vec{0} \in S$ and $\lambda = -1$, then $\forall x \in S, (-1)x + 0 = -x + 0 = -x \in S$

$\square$

[a]Because S is closed under $\cdot, +$

Remark: Typically when X, Y are subspaces in $\mathbb{R}^2$, $X \cup Y$ is not a subspace. Suppose $X \neq Y \wedge X \cup Y \neq X \vee Y$, then take $x \in X \setminus Y$ and $y \in Y \setminus X$. We can see that $x + y \in \mathbb{R}^2$ but $x + y \notin X \cup Y$ since $x \notin Y$ and $y \notin X$. In plain English, $X \cup Y$ is not closed under vector addition.

Definition: The sum of subspaces X and Y is defined as $X + Y \equiv \{x + y \mid x \in X,\ y \in Y\}$ and it is a subspace.

Definition: Let $(V, +, \cdot)$ be a linear space. $x \in V$ is a **Linear Combination** of $\{x_1, x_2, \ldots, x_k\} \subseteq V$ if $\exists \lambda \in \mathbb{R}^k$ such that $x = \sum_i \lambda_i x_i$.

Definition (Span): Let $(V, +, \cdot)$ be a linear space. $A \equiv \{x_1, x_2, \ldots, x_k\} \subseteq V$, then the **Span** of A is:

$$span(A) \equiv \{x \in V \mid x = \sum_{i=1}^{k} \lambda_i x_i,\ \lambda \in \mathbb{R}^k\}$$

Remark: The span of a subset $S \subseteq V$ in a linear space is the smallest linear subspace of V that contains S.

Definition: Let $(V, +, \cdot)$ be a linear space. A subset of vectors $\{x_1, x_2, \ldots, x_k\} \subseteq V$ is **Linearly Independent** if

$$\sum_{i=1}^{k} \lambda_i x_i = 0 \quad \Rightarrow \quad \lambda_i = 0, \forall\, i \in \{1, \ldots, k\}$$

Definition: Let $(V, +, \cdot)$ be a linear space. An infinite set of vectors $X \equiv \{x_1, \ldots\} \subset V$ is **Linearly Dependent** if there exists a finite subset $S \subseteq X$ that is linearly independent.

Definition (Basis): Let $(V, +, \cdot)$ be a linear space. $\forall Z \subseteq V$, a **Basis** of Z is a set W of linearly independent vectors such that $span(W) = Z$.

Remark: Note that most subsets of V do not have a basis. In fact, only subsets that coincide with the span of something (so only subsets that are subspaces) have a basis.

Definition (Dimension): Let $(V, +, \cdot)$ be a linear space and W be a basis of V, the **Dimension** of V is the number of vectors in W (i.e., $dim(V) = |W|$).

Definition: Let $(V, +, \cdot)$ be a linear space. $x \in V$ is a **Non-Negative Linear Combination** of a set of vectors $\{x_1, \ldots, x_k\} \subseteq V$ if $\exists \lambda \in \mathbb{R}_+^{\mathbf{k}}$ such that $x = \sum\limits_{i=1}^{k} \lambda_i x_i$.
The **non-negative cone** of $\{x_1, \ldots, x_k\} \subseteq V$ is the set of all of its non-negative linear combinations.

Definition: Let $(V, +, \cdot)$ be a linear space. The vector $x \in V$ is an **Affine combination** of a set of vector $\{x_1, \ldots, x_k\} \subseteq V$ if $\exists \lambda \in \mathbb{R}^{\mathbf{k}}$ such that $x = \sum\limits_{i=1}^{k} \lambda_i x_i$ AND $\sum\limits_{i=1}^{k} \lambda_i = 1$.
An **affine set** is a set of vectors that are closed under affine combinations.

Definition: Let $(V, +, \cdot)$ be a linear space. Take $x^0 \in V$ and $S \subseteq V$ where S is a subspace of V. A set $A \subseteq V$ is an **Affine subspace** of V **parallel** to S if $A = x^0 + S = \{x \in V \mid x = x^0 + y,\, y \in S\}$. *Note that an affine subspace is not a linear subspace. It is an affine set parallel to the linear space.*

Definition: Let $(V, +, \cdot)$ be a linear space. $x \in V$ is a **Convex Combination** of $\{x_1, \ldots, x_k\} \subseteq V$ if $\exists \lambda \in \mathbb{R}_+^k$ such that $x = \sum\limits_{i=1}^{k} \lambda_i x_i$ and $\sum\limits_{i=1}^{k} \lambda_i = 1$
A convex set is a set that is closed under convex combinations.
Note that this is different from affine sets because we restrict $\lambda \in \mathbb{R}_+^k$ and not $\mathbb{R}^k$

Definition: $\forall n \in \mathbb{N}, p \in \mathbb{R}^n \setminus \{\vec{0}\}$ and $\lambda \in \mathbb{R}$. The **Hyperplane** in $\mathbb{R}^n$ with "normal" p and "level" λ is defined as:

$$H(p, \lambda) \equiv \{x \in \mathbb{R}^n \mid \sum_{i=1}^{k} p_k x_k = \lambda\}$$

Definition: A hyperplane $H(p, \lambda)$ is an affine subspace (linear subspace if and only if $\lambda = 0$)

Definition: A hyperplane $H(p, \lambda)$ divides $\mathbb{R}^n$ into two **Halfspaces**:

$$H_{\geq}(p, \lambda) \equiv \{x \in \mathbb{R}^n \mid \sum_{i=1}^{k} \geq \lambda\}$$
$$H_{\leq}(p, \lambda) \equiv \{x \in \mathbb{R}^n \mid \sum_{i=1}^{k} \leq \lambda\}$$

Definition: Two sets $X, Y \subseteq \mathbb{R}^n$ are **Weakly Separated** by $H(p, \lambda)$ if they lie on opposite sides of $H(p, \lambda)$.

Theorem 3.2: Weak Separating Hyperplane Theorem

If $X, Y \subseteq \mathbb{R}^n \setminus \varnothing$ are *convex* sets and $X \cap Y = \varnothing$, then $\exists p \in \mathbb{R}^n \setminus \{\vec{0}\}$ and $\lambda \in \mathbb{R}$ such that $H(p, \lambda)$ weakly separates X and Y.
i.e., two convex and disjoint sets are on separate sides of some hyperplane.

Theorem 3.3: Separating Hyperplane Theorem

If $X, Y \subseteq \mathbb{R}^n \setminus \varnothing$ are *convex* sets, X is *open*, and $X \cap Y = \varnothing$, then $\exists p \in \mathbb{R}^n \setminus \{\vec{0}\}$ and $\lambda \in \mathbb{R}$ such that $\sum_{i=1}^{k} p_k x_k < \lambda \leq \sum_{i=1}^{k} p_i y_i, \forall x \in X,\ y \in Y$.

Theorem 3.4: Supporting Hyperplane Theorem

Assume that $X \subseteq \mathbb{R}^n \setminus \varnothing$ is *convex and closed* and $int(X) \neq \varnothing$, then $\forall b \in bd(X)$, $\exists\, p \in \mathbb{R}^n \setminus \varnothing$ such that $\sum_{i=1}^{k} p_k b_k \leq \sum_{i=1}^{k} p_k x_k$, $\forall x \in X$ and $\sum_{i=1}^{k} p_k b_k < \sum_{i=1}^{k} p_k x_k$, $\forall x \in int(X)$

Definition: $X, Y \subseteq \mathbb{R}^n$ are **Strongly Separated** by $H(p, \lambda)$ is $\exists \varepsilon \in \mathbb{R}_{++}$ such that $\forall x \in X,\ y \in Y,\ \sum_{i=1}^{k} p_k x_k \leq \lambda - \varepsilon < \lambda < \lambda + \varepsilon \leq \sum_{i=1}^{k} p_k y_k$

Theorem 3.5: Strong Separating Hyperplane Theorem

If $X, Y \subseteq \mathbb{R}^n \setminus \varnothing$ are *convex*, X is *compact*, Y is *closed*, and $X \cap Y = \varnothing$, then $\exists p \in \mathbb{R}^n \setminus \{\vec{0}\}$ such that $H(p, \lambda)$ strongly separates X and Y

Proposition: Let $C \subseteq \mathbb{R}^n \setminus \varnothing$ be closed and convex. Then the intersection of all the closed halfspaces that contain C is equal to C.

Proof:

$\forall a \notin C$, $\{a\}$ and C are disjoint. $\{a\}$ is trivially compact and C is assumed to be a closed set. So by the strong separating hyperplane theorem, $\exists H(p, \lambda)$ that strongly separates $\{a\}$ and C. Hence at least one of the halfspaces contains C and not $\{a\}$. So the intersection of halfspaces containing C will not contain $\{a\}$. $\square$

Theorem 3.6: Farka's Lemma

Let $m, n \in \mathbb{N}$, $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$. Then either
$\exists x \in \mathbb{R}_+^m$ such that $Ax = b$
or
$\exists y \in \mathbb{R}_+^m$ such that $A^T y \in \mathbb{R}_+^m$ and $y^T b < 0$

- m different n-dimensional vectors
- x is a vector of weights

3.2 Norm and Inner Products

Definition (Norm): A **Norm** on a linear space is a function $\|\cdot\| : V \to \mathbb{R}_+$ that satisfies $\forall x, y \in (V, +, \cdot)$:

(i) $\|x\| = 0$ if and only if $x = \vec{0}$

(ii) $\forall a \in \mathbb{R}$, $\|ax\| = \|a\|\|x\|$

(iii) (Triangle Inequality) $\|x + y\| \leq \|x\| + \|y\|$

The standard **p-norm** on $\mathbb{R}^n$ is $\|\cdot\|_p = d_p(x, \vec{0})$, $p \in [1, \infty)$.
In $\mathbf{C}[a, b]$, the L_p-norm is $\left(\int_a^b |f(t)|^p dt\right)^{\frac{1}{p}}$.
In $\mathbf{C}[a, b]$, the L_∞-norm is $\|f\|_\infty = \max_{t \in [a,b]} |f(t)|$

Notice that norm only gives us the magnitude of a vector, but not the direction of the vector. It might not seem significant but the directions of vectors can actually give us a lot of useful information. In order to preserve the directions of vectors to an extent, we shall use the **inner product** of vectors.

Definition: Let $(V, +, \cdot)$ be a linear space. The **Inner Product** $\langle \cdot, \cdot \rangle : V \times V \to R$ is an operation such that $\forall x, y, z \in V, a, b \in \mathbb{R}, x \neq \vec{0}$:

(i) $\langle \vec{0}, \vec{0} \rangle = 0$ and $\langle x, x \rangle > 0$

(ii) (**Symmetry**) $\langle x, y \rangle = \langle y, x \rangle$

(iii) (**Linearity**) $\langle x, ay + bz \rangle = a\langle x, y \rangle + b\langle x, z \rangle$

Remark: Note that the level of "structured-ness" goes as "$X \to d_X \to \|\cdot\|_X \to \langle \cdot, \cdot \rangle_X$". So a basic set (a simple collection of objects) has the least structure, and a space with an inner product has the most structure.
Dot product $(\cdot)$ is a special case of inner product. In $\mathbb{R}^n$, $x \cdot y = \sum_{i=1}^{n} x_i y_i$. With dot product as the inner product, we can re-define the norm in $\mathbb{R}^n$ to be:

$$\|x\| \equiv \sqrt{\langle x, x \rangle} = \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2} \equiv \|x\|$$

and we can redefine the metric on $\mathbb{R}^n$:

$$d_X(x, y) = \|x - y\| = \langle x - y, x - y \rangle = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \equiv d_2(x, y)$$

Definition: Let $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ be a linear space. We say $x, y \in V$ are orthogonal $(\perp)$ to each other if $\langle x, y \rangle = x \cdot y = 0$.

Theorem 3.7: Dot Product Preserves Directions

$$\forall x, y \in \mathbb{R}^n, x \cdot y = \underbrace{\|x\|\|y\|}_{\text{Magnitude}} \cdot \underbrace{cos(\theta)}_{\text{Angle between } x \text{ and } y}$$

So if $x \perp y, cos(\theta) = cos(\frac{\pi}{2}) = 0 \;\Rightarrow\; x \cdot y = \|x\|\|y\| \cdot 0 = 0$

Definition: A set of vectors $X \subseteq \mathbb{R}^n$ is **Orthonormal** if X is orthogonal (every vector is orthogonal to the other vectors) AND $\forall x \in X$, $|X| = 1$.

Theorem 3.8: Cauchy-Schwartz Inequatlity

Let $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ be a linear space. $\forall x, y \in V$,

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle$$

Proof 3.8: Cauchy-Schwartz Inequatlity

For this proof, we want to show that $\forall x, y \in V$, $\langle x, y\rangle \cdot \langle x, y\rangle \leq \langle x, x\rangle \cdot \langle y, y\rangle$. Since inner products have a special property with $\vec{0}$, we will discuss the case of $\vec{0}$ separately.

Case 1:
WLOG, let $y = \vec{0}$, we can actually show that the inequality is equality in this case.

$$\langle x, \vec{0}\rangle = \underbrace{\langle x, \vec{0} + \vec{0}\rangle = \langle x, \vec{0}\rangle + \langle x, \vec{0}\rangle}_{\text{Linearity of inner products}}$$

Since $\langle x, \vec{0}\rangle = \langle x, \vec{0}\rangle + \langle x, \vec{0}\rangle$, it must be that $\langle x, \vec{0}\rangle = 0$. Since $y = \vec{0}$, $\langle y, y\rangle = 0$, so we have:

$$\langle x, \vec{0}\rangle \cdot \langle x, \vec{0}\rangle = 0 \cdot 0 = \langle x, x\rangle \cdot 0$$

Case 2:
Take $x \neq \vec{0}$, $y \neq \vec{0}$, then we know that $\langle x, y\rangle \neq 0$. We will then define

$$z \equiv x - \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot \underbrace{y}_{=0}$$

Then we have that

$$\langle y, z\rangle = \langle y, \underbrace{x - \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y}_{=z}\rangle = \underbrace{\langle y, x\rangle - \langle y, \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y\rangle}_{\text{By linearity}}$$
$$= \langle x, y\rangle - \frac{\langle x, y\rangle}{\cancel{\langle y, y\rangle}} \cdot \cancel{\langle y, y\rangle} = \langle x, y\rangle - \langle x, y\rangle = 0$$

Rearranging this, we get

$$\begin{aligned}
\langle x, x\rangle &= \langle z + \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y, z + \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y\rangle \\
&= \langle z, z + \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y\rangle + \langle \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y, z + \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y\rangle \\
&= \langle z, z\rangle + \langle z, \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y\rangle + \langle \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y, z\rangle + \langle \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y, \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot y\rangle
\end{aligned}$$

$$= \langle z, z\rangle + \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot \underbrace{\langle z, y\rangle}_{=0} + \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot \underbrace{\langle y, z\rangle}_{=0} + \frac{\langle x, y\rangle}{\langle y, y\rangle} \cdot \frac{\langle x, y\rangle}{\cancel{\langle y, y\rangle}} \cdot \cancel{\langle y, y\rangle}$$

$$= \underbrace{\langle z, z\rangle}_{>0} + \frac{\langle x, y\rangle \cdot \langle x, y\rangle}{\langle y, y\rangle} \geq \frac{\langle x, y\rangle \cdot \langle x, y\rangle}{\langle y, y\rangle}$$

Multiplying both sides of the inequality by $\langle y, y\rangle$, we get

$$\langle x, x\rangle \geq \frac{\langle x, y\rangle \cdot \langle x, y\rangle}{\langle y, y\rangle} \Rightarrow \langle x, x\rangle \cdot \langle y, y\rangle \geq \frac{\langle x, y\rangle \cdot \langle x, y\rangle}{\cancel{\langle y, y\rangle}} \cdot \cancel{\langle y, y\rangle}$$

Rearranging the equation, we get the Cauchy-Schwartz Inequality.

$$\langle x, y\rangle \cdot \langle x, y\rangle \leq \langle x, x\rangle \cdot \langle y, y\rangle$$

□

Definition: A linear space endowed with an inner product is called a **pre-Hilbert space**. (e.g., $(V, +, \cdot, \langle \cdot, \cdot\rangle)$)

Definition: A **Hilbert space** is a pre-Hilbert space that is also endowed with a *norm* ($\|\cdot\|$) and a *metric* (d_X), and it is *complete* with respect to d. (e.g., $(V, +, \cdot, d_X, \|\cdot\|_X, \langle \cdot, \cdot\rangle)$)

Definition: A **Banach space** is a linear space endowed with a *norm* and a *metric*, and is *complete* (no requirement of inner product like the Hilbert space). (e.g., $(V, +, \cdot, d_X, \|\cdot\|_X)$)

Theorem 3.9: Projection Theorem

Let H be a Hilbert space and take $x \in H$. Let $M \subseteq H$ be a linear subspace of H. Then

$$\exists m_x \in M \;\; s.t. \; \forall m \in M, \;\; \|x - m_x\| \leq \|x - m\|$$

Also observe that $(x - m_x) \perp M$

Definition: Let $(X, +, \cdot)$ and $(Y, \oplus, \odot)$ be two linear spaces. A function $f : X \to Y$ is said to be **Linear** if $\forall x, z \in X$, $\forall \alpha \in \mathbb{R}$:

(i) $f(x + z) = f(x) \oplus f(z)$

(ii) $f(\alpha x) = \alpha \odot f(x)$

Specifically, for $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(x) = Ax$ where A is a $m \times n$ matrix.

Definition: Let $(X, +, \cdot)$ and $(V, \oplus, \odot)$ be two linear spaces. We say these two spaces are **Isomorphic** if there exists an *invertible linear function* $f : V \to X$. In this case, f is called an **Isomorphism**.

Important Proposition: Two linear spaces over the same field are *isomorphic* if and only if they have the same dimension.

Definition: Let $(X, +, \cdot)$ and $(Y, \oplus, \odot)$ be two linear spaces. A function $f : X \to Y$ is **Affine** if

$$\forall \alpha \in \mathbb{R},\ x, z \in X,\ f(\alpha x + (1-\alpha)z) = \alpha \odot f(x) \oplus (1-\alpha) \odot f(z)$$

Definition: Let $(V, +, \cdot)$ be a linear space. The constant $\lambda \in \mathbb{R}$ is called an **Eigenvalue**[12] of the square matrix $A_{n \times n} \in \mathbb{R}^n \times \mathbb{R}^n$ if

$$\exists v \in \mathbb{R}^n \setminus \vec{0},\ Av = \lambda v$$

Equivalent Definition: Let $(V, +, \cdot)$ be a linear space. The constant $\lambda \in \mathbb{R}$ is called an **Eigenvalue** of the square matrix $A_{n \times n} \in \mathbb{R}^n \times \mathbb{R}^n$ if

$$det(A - \lambda I_{n \times n}) = 0$$

Remark: To calculate the eigenvalue, one can solve for the polynomial characterized by $det(A - \lambda I) = 0$. This polynomial is called the **characteristic equation**.

Definition: A square matrix $A_{n \times n} \in \mathbb{R}^n \times \mathbb{R}^n$ is said to be **Singular** if it is *not invertible* ($det(A) = 0$).

Definition: An eigenvector $v \in V$ is a vector such that $Av = \lambda v$. i.e., the eigenvectors are always paired with eigenvalues.

[12] See Simon and Blume Ch.23 for a discussion of eigenvalues and eigenvectors.

4 Optimization

By the Taylor theorem, real analytic functions can be **locally approximated** by polynomials. So it would make sense for us to study (local) optimization of polynomials in order to take a peek into how certain functions behave. Specifically, we can represent polynomials in quadratic forms.

Definition: A quadratic function on X is a function of the form $Q(x) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_1x_2 + a_{22}x_2^2 = x^TAx$ where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

A matrix A is	**positive definite**	if	$\forall x \neq 0,\ x^TAx > 0$
A matrix A is	**positive semi-definite**	if	$\forall x \neq 0,\ x^TAx \geq 0$
A matrix A is	**negative definite**	if	$\forall x \neq 0,\ x^TAx < 0$
A matrix A is	**negative semi-definite**	if	$\forall x \neq 0,\ x^TAx \leq 0$
A matrix A is	**indefinite**	if	$\exists\, x, \tilde{x} \in X \setminus \vec{0},\ x^TAx > 0 \wedge \tilde{x}^TA\tilde{x} < 0$

4.1 Unconstrained Optimization

Definition: Take the function $f : \mathbb{R}^m \to \mathbb{R}$ and $x \in \mathbb{R}^m, h \in \mathbb{R}$, then $\forall k \in \{1, \ldots, m\}$, the k^{th} **partial derivative** of f is:

$$f_{x_k} = \frac{\partial f}{\partial x_k} = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{k-1}, x_k + h, x_{k+1} \ldots, x_m) - f(x_1, \ldots, x_{k-1}, x_k, x_{k+1} \ldots, x_m)}{h}$$

Definition: Take the function $f : \mathbb{R}^m \to \mathbb{R}$ and $x \in \mathbb{R}^m$. The **Gradient** at x (denoted $\nabla f(x) : \mathbb{R}^m \to \mathbb{R}^m$) and the **Derivative** (denoted $Df(x)$) at x are:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_2} \end{pmatrix} \text{ and } Df(x) = (\nabla f(x))^T = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m} \right)$$

Definition: For all open sets $O \subseteq \mathbb{R}^n$, f is **Differentiable** on O if $Df(x)$ is well-defined for all $x \in O$.

Definition: Take $X \subseteq \mathbb{R}^m$, $x^* \in int(X)$ is called a **Critical Point** if $Df(x^*) = \vec{0} \in \mathbb{R}^m$.

Definition: Take the function $f : \mathbb{R}^m \to \mathbb{R}$. The **Cross-Partial** (derivative) denoted as $\frac{\partial^2 f}{\partial x_i \partial x_j} : \mathbb{R}^m \to \mathbb{R}$ is the j^{th} partial derivative of $\frac{\partial f}{\partial x_i}$. i.e., $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})$

Definition: The Hessian matrix (H) of $f : \mathbb{R}^m \to \mathbb{R}$ is:

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & & \cdots & \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m} \end{pmatrix}$$

Why do we use the Hessian matrix?

Take $x, h \in \mathbb{R}^m$, $f : \mathbb{R}^m \to \mathbb{R}$, we can approximate f locally at $x + h$ with:

$$f(x+h) = f(x) + \frac{\partial f}{\partial x_1}(x)h_1 + \cdots + \frac{\partial f}{\partial x_m}(x)h_m + \underbrace{R(x,h)}_{\text{The remaining term of the approximation}}$$

By the Taylor theorem, we know that $\frac{R(x,h)}{\|h\|} \xrightarrow{h \to 0} 0$. We can also further approximate with the second order Taylor polynomial using:

$$f(x^* + h) = f(x^*) + Df(x^*)h + \frac{1}{2}h^T H(x^*)h + R_2(X^*, h)$$

Since x^* is a critical point, we know that $Df(X^*) = 0$. We also know that, locally, $R_2(x^*, h)$ is negligibly small and hence

$$f(x^* + h) - f(x^*) \approx \frac{1}{2}h^T H(x^*)h$$

So the local derivative of f at x^* can be approximated by the definiteness of the Hessian matrix $H(x^*)$

Proposition: Take $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}$, where f is twice-continuously-differentiable (C^2) and $x^8 \in \int(X)$ is a critical point, then

- If $H(x)$ is Positive Definite, then x^* is a **strict local minimum**
- If $H(x)$ is Negative Definite, then x^* is a **strict local maximum**
- If $H(x)$ is Indefinite, then x^* is **neither**

4.2 Convex Optimization

Definition: The function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **Strictly Quasi-Concave** on (a, b) if

$$\forall x \in (a, b) \subseteq \mathbb{R}^n,\ f(x) > min(f(a), f(b))$$

Proposition: Let $X \subseteq \mathbb{R}^n$ be a convex set and $f : X \to \mathbb{R}$ be C^2, then

If H(x) is Positive Definite, then $\forall x \in X$, f is **globally convex**

If H(x) is Negative Definite, then $\forall x \in X$, f is **globally concave**

Global convexity/concavity implies unique optimum. Our usual consumer problem:

$$\max_{x \in \mathbb{R}^n} f(x)\ s.t.\ g(x) \leq b \in \mathbb{R}$$

So we can typically simplify our problem to

(1) If f is continuous and $\{x \in \mathbb{R}^n \mid g(x) \leq b\}$ is compact, then we know a solution exists by the Weirestrass Maximum Theorem.

(2) If f is continuous AND strictly quasi-concave and $\{x \in \mathbb{R}^n \mid g(x) \leq b\}$ is compact and convex, then we know that the solution is unique.

Theorem 4.1: Kuhn-Tucker Theorem for Inequality Constraints

Take $m, n \in \mathbb{N}$, $f : \mathbb{R}^n \to \mathbb{R}$, f is C^1, and $\forall k \in \{1, \ldots, n\}$, let $g_k : \mathbb{R}^n \to \mathbb{R}$ be C^1

- Assume that x^* is a local maximizer of f on the constraint set defined by $g(x) \leq b \in \mathbb{R}^m$
- Assume that constraints 1 through l are binding
- (NDCQ) Suppose that $Dg_1(x^*), Dg_2(x^*), \ldots, Dg_l(x^*)$ are linearly independent

Define the Lagrangian $\mathcal{L}(x, \lambda) \equiv f(x) - \sum_{k=1}^{m} \lambda_k (g_k(x) - b_k)$, then $\exists \lambda^* \in \mathbb{R}^m_+$ such that

(i) $\frac{\partial \mathcal{L}}{\partial x_k} = 0$

(ii) (Complementary slackness condition) $\forall k \in \{1, \ldots, m\}$, $\lambda^* \cdot (g_k(x^*) - b) = 0$, $\lambda^* \geq 0$

(iii) $\forall k \in \{1, \ldots, m\}$, $g_k(x^*) \leq b_k$

Definition: The **Jacobian** of a system of m equations $g_1, g_2, \dots, g_m$ of n variables $x_1, x_2, \dots, x_n$ is the matrix:

$$J = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

Remarks: The intuition behind the Kuhn-Tucker theorem is that $Df(x^*) = (\lambda^*)^T J(g(x^*))$. i.e., the gradient of the objective function is a linear combination of the gradients of the constraints.

Theorem 4.2: Kuhn-Tucker Theorem with Non-Negativity Constraints

Take $m, n \in \mathbb{N}$, $f : \mathbb{R}^n \to \mathbb{R}$, f is C^1, and $\forall k \in \{1, \dots, n\}$, let $g_k : \mathbb{R}^n \to \mathbb{R}$ be $C61$ as well

- Assume that x^* is a local maximizer of f on the constraint set defined by $g(x) \leq b \in \mathbb{R}^m$ *AND* $\mathbf{0} \leq \boldsymbol{x}$
- Assume that constraints 1 through l are binding
- (NDCQ) Suppose that $Dg_1(x^*), Dg_2(x^*), \dots, Dg_l(x^*)$ are linearly independent

Let $\mathcal{L}(x, \lambda) = f(x) - \sum_{k=1}^{m} \lambda_k (g_k(x) - b_k)$, then $\exists \lambda^* \in \mathbb{R}^m_+$ such that

(i) $\frac{\partial \mathcal{L}}{\partial x_k} = 0$

(ii) (Complementary slackness condition) $\forall k \in \{1, \dots, m\}$, $\lambda^*(g_k(x^*) - b) = 0$

(iii) $\forall k \in \{1, \dots, m\}$, $g_k(x^*) \leq b_k$

(iv) $x_i \cdot \frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda^*) = 0$

4.3 Comparative Statics

Theorem 4.3: Implicit Function Theorem on $\mathbb{R}^2$

Let $G(x, y)$ be a C^1 function[a] on $N_\varepsilon^{d_2}((x_0, y_0))$ and $G(x_0, y_0) = c \in \mathbb{R}$. If $\frac{\partial F(x_0,y_0)}{\partial y} \neq 0$, then $\exists \delta \in \mathbb{R}_{++}$ and a C^1 function $y(x)$ defined on $I \equiv (x_0 - \delta, x_0 + \delta)$ such that

(i) $\forall x \in I$, $G(x, y(x)) = c$

(ii) $y(x_0) = y_0$

(iii) $\dfrac{dy}{dx}(x_0) = -\dfrac{\partial G(x_0, y_0)/\partial x}{\partial G(x_0, y_0)/\partial y}$

[a] C^n means a $n-$times continuously differentiable function

Example: Consider $G(x, y) = x^2 - 3xy + y^3 - 7 = 0$ and $(4, 3)$ is a solution. Where x is dollars invested in advertisements and y is dollars in sales. How would sales be impacted by an increase in advertising expenditure? More precisely, what is $\frac{dy}{dx}$ at $x = 4$? We take the partial derivative of $G(x, y)$ with respect to x and y and evaluate at the point (4,3).

$$\frac{\partial G}{\partial y}|_{(4,3)} = -3x + 3y^2|_{(4,3)} = 27 - 12 = 15$$
$$\frac{dy}{dx} = -\frac{2x - 3y}{-3x + 3y^2}|_{(4,3)} = -\frac{8 - 9}{15} = \frac{1}{15}$$

This means for every additional dollar spent on advertising, sales increase by 15 dollars.

Theorem 4.4: Implicit Function Theorem on $\mathbb{R}^m$

Let $G : \mathbb{R}^{m+1} \to \mathbb{R}$ be a C^1 function on $N_\varepsilon^{d_2}((x_1^*, x_2^*, \ldots, y^*))$ and $G(x^*, y^*) = c \in \mathbb{R}$. If $\frac{\partial G(x_0,y_0)}{\partial y} \neq 0$, then $\exists \delta \in \mathbb{R}_{++}$ and a C^1 function $y(x)$ defined on $I \equiv (x_0 - \delta, x_0 + \delta)$ such that

(i) $\forall x \in N_\delta^{d_2}(x^*)$, $G(x, y(x)) = c$

(ii) $y(x^*) = y^*$

(iii) $\dfrac{dy}{dx}(x^*) = -\dfrac{\partial G(x^*, y^*)/\partial x}{\partial G(x^*, y^*)/\partial y}$

Theorem 4.5: Envelope Theorem (Unconstrained)

Let $f(x,a)$ be a C^1 function on $x \in \mathbb{R}^n$ and $a \in \mathbb{R}$. Suppose that

$$x^*(a) = arg\,\underset{x}{max}\, f(x,a)$$

is a C^1 function of a, then

$$\frac{d}{da} f(x^*(a),a) = \frac{\partial}{\partial a} f(x^*(a),a)$$

In other words, on the margin, the total derivative is equal to the partial derivative because the changes in the maximizer are too small to pass through to impact the objective function.

Motivating Economic Example:

Now let's consider the classic economist's problem again. Say we need to maximize a utility function $u(p,x)$ subject to our budget constraint $B(p) \in X$ where X is the set of consumption bundles and P is the set of prices. Notice that in this kind of problem, the price p is a parameter of both our objective function and our constraint set. Let us assume that $B(p)$ is compact, so that we can try to get some nice results. Formally, our problem is:

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

By the Weierstrass Theorem of Maximum, we know that if $u(p,x)$ is continuous on x in $B(p)$, we can find a maximizer x^* within $B(p)$, $\forall p \in P$, such that $x(p) \equiv \{\hat{x} \in B(p) \mid u(p,\hat{x}) = V(p)\}$ where $V(p) \equiv \max_{x \in B(p)} u(p,x)$. Moreover, we can actually show that IF $u(p,x)$ is also continuous in p, then $V(p)$ is continuous in p and, $X(p)$, the correspondence for utility maximizing bundles, is upper-hemi-continuous.

Theorem 4.6: Berge's Theorem of Maximum

- Let (P, d_P), (x, d_X) be metric spaces.
- $u : P \times X \to \mathbb{R}$ is continuous on $P \times X$.
- $B : P \rightrightarrows X$ is compact-valued and continuous on P.

If the above assumptions hold, then we have the following results

⋆ $V : P \to \mathbb{R}$, the value function defined as $V(p) \equiv \max\limits_{x \in B(p)} u(p,x)$, is continuous on P.

⋆ $X : P \rightrightarrows X$, the maximizer correspondence defined as:
$x(p) \equiv \{x \in B(p) \mid u(p,x) = V(p)\} = \operatorname*{argmax}\limits_{x \in B(p)} u(x,p)$, is an upper-hemi-continuous and compact-valued correspondence.

Proof 4.6: Berge's Theorem of Maximum (Special Case)

We will prove this for the special[a] case in $\mathbb{R}^n$.

Step 1: We want to show that $x(p)$ is well-defined.
Since $B(p)$ is compact-valued, $\forall p \in P$, $B(p)$ is compact. Since u is continuous on P, we know that u attains a maximum $x(p) \in B(p)$ $\forall p \in P$.

Step 2: We want to show that $x : P \rightrightarrows X$ is **compact-valued**.

(a) (Boundedness) Since $B(p)$ is compact, it is bounded. Since $x(p)$ is a subset of a bounded set, $x(p)$ is bounded.

(b) (Closedness) Take a sequence $(x_k)_{k=1}^{\infty} \in x(p)$ such that $x_k \to \hat{x}$. Since $(x_k)_{k=1}^{\infty} \in x(p) \subseteq B(p)$ and $B(p)$ is closed, we know that $\hat{x} \in B(p)$. Now we need to show that $\hat{x}$ is actually in $x(p)$.

Suppose otherwise that $\hat{x} \in B(p) \backslash x(p)$ for a given p, then we know that $\exists y \in B(p)$ such that $u(p,y) > u(p,\hat{x})$. Since the inequality is strict, we know that $\exists \delta \in \mathbb{R}_{++}$ such that $\forall w \in N_{\delta}^{d_2}(\hat{x})$, $u(p,y) > u(p,w)$. Since $x_k \to \hat{x}$ we know that $\exists M \in \mathbb{N}$ such that $\forall k > M$, $x_k \in N_{\delta}^{d_2}(\hat{x})$. This means that $\forall k > M, x_k \notin x(p)$. But the sequence is constructed so that $\forall k \in \mathbb{N}$, $x_k \in x(p)$. Hence by contradiction,

$\hat{x} \in x(p)$ and hence $x(p)$ is closed.

Since $x(p)$ is closed for any given $p \in P$, we know that $x(p)$ is compact/compact-valued.

Step 3: We want to show that $x : P \rightrightarrows X$ is **upper-hemi-continuous**. We can achieve this by showing that X is compact AND $x(p)$ has a closed graph.
Take the sequences $(p_k)_{k=1}^{\infty}$ and $(x_k)_{k=1}^{\infty} \in x(p_k)$ such that $p_k \to \bar{p}$ and $x_k \to \bar{k}$.

(a) Since $B : P \to X$ is assumed to be upper-hemi-continuous (and lower-hemi-continuous, but that doesn't matter here) and is closed-valued, we know that $B(p)$ has a closed graph in $P \times X$ and $\bar{x} \in B(p)$. So we just need to show that $\bar{x} \in x(\bar{p})$.
Suppose otherwise that $\bar{x} \notin x(\bar{p})$, then $\exists y \in B(p)$ such that $u(\bar{p}, y) > u(\bar{p}, \bar{x})$. Since $B : P \to X$ is continuous, it is lower-hemi-continuous. Then $\exists (y_k)_{k=1}^{\infty} \to y$ such that $y_k \in B(\bar{p})$ and $(p_k, y_k) \to (\bar{p}, y)$. This means that $\exists M \in \mathbb{N}$ such that $\forall k > M$, $u(p_k, y_k) > u(p_k, x_k)$. But this means that $(x_k)_{k=1}^{\infty} \notin x(\bar{p})$. So by contradiction, $\bar{x} \in x(\bar{p})$ and we know that $x(p)$ has a closed graph.

(b) Since $x(p)$ has a closed graph and we have shown that X is compact in step 2, we know that $x : P \rightrightarrows X$ is upper-hemi-continuous.

Step 4: We want to show that $V(p) \equiv \max_{x \in B(p)} u(p, x)$ is continuous on P
Notice that we can rewrite $V(p)$ as $u(p, x(p))$, so what we need to show is that $\forall p_0 \in P$, $\forall$ open $O \subseteq U \subseteq \mathbb{R}$ $s.t.$ $u(p_0, x(p_0)) \in O$,

$$\exists\, \delta_O \in \mathbb{R}_{++}\, s.t. \forall p \in N_{\delta_O}^{d_P}(p_0),\; u(p, x(p)) \in O$$

Since $u : P \times X \to \mathbb{R}$ is assumed to be continuous on $P \times X$, we know that $\forall p_0 \in P$, $\forall$ open $O \subseteq U \subseteq \mathbb{R}$ $s.t.$ $u(p_0, x)) \in O$. Notice that $x(p)$ is upper-hemi-continuous so $\forall p_0 \in P, \forall O' \in X$ $s.t.$ $x(p_0) \subseteq O', \exists \delta_{O'} \in \mathbb{R}_{++}$ $s.t.$ $x(N_{\delta_{O'}}^{d_P}(p_0)) \subseteq O'$. Now for each $p_0 \in P$, take $\delta^* \equiv \min\{\delta_O, \delta'_O\}$, then $\forall$ open $O \subseteq U \subseteq \mathbb{R}$ $s.t.$ $u(p_0, x(p_0)) \in O$, we have $\forall p \in N_{\delta^*}^{d_P}(p_0) = (N_{\delta^*}^{d_P}(p_0), N_{\delta^*}^{d_P}(x(p_0)))$, $V(p) = u(p, x(p)) \in O$ □

[a] See Efe Ok page 306 for general proof.

Theorem 4.7: Envelope Theorem (Constrained)

Let $f(x,a)$ be a C^1 function on $x \in \mathbb{R}^n$ and $a \in \mathbb{R}$. Suppose that we want to maximize $f(x,a)$ on a compact subset $A \subset \mathbb{R}^n \times \mathbb{R}$ and NDCQ is satisfied everywhere.
The solution $x^*(a)$ can be characterized by the Lagrangian

$$\mathcal{L}(x,a) = f(x,a) - \sum_{i=1}^{m} \lambda_i (g_i(x^*) - b_i)$$

where $\sum_{i=1}^{m} g_i(x^*) \leq b_i$ describe the constraint set A.

Suppose that the solution

$$x^*(a) = \underset{x \in A}{arg\,max}\, f(x,a)$$

is a C^1 function and lies along the binding constraints. Then,

$$\frac{d}{da} f(x^*(a), a) = \frac{\partial \mathcal{L}(x,a)}{\partial a}$$

and λ_i can be interpreted as the *shadow price* of relaxing the ith binding constraint.

In other words, the Lagrange multiplier gives you the marginal change in the objective function when you allow the maximizer to move slightly outside of the constraint set.

4.4 Supermodularity and Monotonic Comparative Statics[13]

Definition: Ordering/Partial Ordering in $\mathbb{R}^n$

For $x, y \in \mathbb{R}^n$ and x_k, y_k represent the k^{th} element of x and y, we have:

$$\begin{aligned}
x = y &\quad \text{if} \quad \forall k \in \{1,\dots,n\},\, x_k = y_k \\
x \geq y &\quad \text{if} \quad \forall k \in \{1,\dots,n\},\, x_k \geq y_k \\
x > y &\quad \text{if} \quad \forall k \in \{1,\dots,n\},\, x_k \geq y_k \wedge \exists j \in \{1,\dots,n\},\, x_j > y_j \\
x >> y &\quad \text{if} \quad \forall k \in \{1,\dots,n\},\, x_k > y_k
\end{aligned}$$

Notice that the ordering $\geq$ is complete in $\mathbb{R}^1$ but not $\mathbb{R}^n$ where $n \in \mathbb{N} \setminus \{1\}$

[13]For more details, consult Sundaram Ch.10

Definition: $\forall x, y \in \mathbb{R}^n$, the **Meet** of x and y is defined as:

$$x \wedge y \equiv (\min\{x_1, y_1\}, \min\{x_2, y_2\}, \ldots, \min\{x_n, y_n\})$$

Definition: $\forall x, y \in \mathbb{R}^n$, the **Join** of x and y is defined as:

$$x \vee y \equiv (\max\{x_1, y_1\}, \max\{x_2, y_2\}, \ldots, \max\{x_n, y_n\})$$

Remark: This means $x \wedge y \leq x, y \leq x \vee y$

Definition: $X \subseteq \mathbb{R}^n$ is a **(compact) sublattice** of $\mathbb{R}^n$ if $\forall x, y \in X$ both $x \wedge y$ and $x \vee y$ are in X (and it is compact in $\mathbb{R}^n$). i.e., X contains all meets and joins of its elements, a "complete grid".

Definition: $x^* \in X \subseteq \mathbb{R}^n$ is the **Greatest Element** if $\forall x \in X$, $x^* \geq x$.

Theorem 4.8

If X is a *non-empty* and *compact sublattice* of $\mathbb{R}^n$, then X has a *greatest element*

Definition: The function $f : X \times Y \to \mathbb{R}$ is said to be **Supermodular** in $X \times Y$ if $\forall z' \equiv (x', y') \in X \times Y$, $z'' \equiv (x'', y'') \in X \times Y$, we have

$$f(z') + f(z'') \leq f(z' \wedge z'') + f(z' \vee z'')$$

and the inequality is strict if z', z'' are not comparable.
Eguia uses the example for illustration purposes: The sum of Milk and cookies separate is less than or equally as good as the sum of 'either' milk or cookies, and 'both' together.

Definition: The function $f : X \times Y \to \mathbb{R}$ is said to be **Supermodular in X for any $y \in Y$** if $\forall y \in Y$, $x, x' \in X$, we have

$$f(x, y) + f(x', y) \leq f(x \wedge x', y) + f(x \vee x', y)$$

Definition: The function $f : X \times Y \to \mathbb{R}$ is said to be **Increasing in Differences in**

X* × *Y if $\forall (x, y), (x', y') \in X \times Y$ such that $x' \geq x$ and $y' \geq y$ we have

$$f(x', y') - f(x', y) \geq f(x, y') - f(x, y)$$

Theorem 4.9

Take $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$. Suppose $f : X \times Y \to \mathbb{R}$ is supermodular in X for all $y \in Y$ and f is increasing in differences in $X \times Y$, then

$$\text{Increasing differences in any two of } m+n \text{ components of } X \times Y$$
$$\Rightarrow\ f \text{ is supermodular in } X \times Y$$

Theorem 4.10

Take Z as an open sublattice of $\mathbb{R}^n$. A C^2 function $f : Z \to \mathbb{R}$ is supermodular on Z if and only if

$$\forall\, i, j \in \{1, \ldots, m\}, z \in Z,\ \frac{\partial^2 f}{\partial z_i \partial z_j} \geq 0$$

Theorem 4.11: (From Topkis 1979)

Take a compact subset of $\mathbb{R}$, $X \subseteq \mathbb{R}$ and T a partially ordered set. Assume that $f : X \times T]conv\mathbb{R}$ has increasing differences in (x, t) and is continuous on X. Then $x^*(t) = \underset{x \in X}{\operatorname{argmax}}\, f(x, t)$ exists and has a maximum $x^*_+(t)$ and $x^*_-(t)$ (There exists a highest and lowest maximizer). Moreover, $\forall t, t' \in T\ s.t.\, t' \geq t$, we have

$$x^*_+(t') \geq x^*_+(t),\ x^*_-(t') \geq x^*_-(t)$$

Theorem 4.12: (See Efe Ok page 276 for proof)

Take $X \subseteq \mathbb{R}^n$ to be a compact sublattice of $\mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ to be a sublattice. Let $f : X \times Y \to \mathbb{R}$ be continuous on X, $\forall y \in Y$ and define $x^*(y) \equiv \underset{x \in X}{\operatorname{argmax}}\, f(x, y)$. If f satisfies increasing differences in $X \times Y$ and supermodular in X, $\forall y \in Y$, then $\forall y' \in Y\ s.t.\ y' > y$ and $x \in x^*(y)$, $x' \in x^*(y')$ implies $x' \geq x$.

4.5 Fixed Point Theorems

Definition (Fixed Point): Let X be a non-empty set and $g : X \rightrightarrows X$ a self-mapping correspondence. We say that a point $x \in X$ is a fixed point of g if $x \in g(x)$. If g is a self-mapping function instead, $x \in X$ is a **Fixed Point** if $g(x) = x$.

Theorem 4.13: Tarski's Fixed Point Theorem

$\forall n \in \mathbb{N}, \forall f : [0,1]^n \to [0,1]^n$

$$f_k : [0,1]^n \to [0,1] \text{ is non-decreasing} \;\Rightarrow\; f \text{ attains a fixed point}$$

Theorem 4.14: Brouwer's Fixed Point Theorem

Let $x \subseteq \mathbb{R}^n$ be a non-empty, compact, and convex subset of $\mathbb{R}^n$. Any continuous self-mapping $f : X \to X$ attains a fixed point in X.

Proof 4.14: Baby Brouwer

We will prove Brouwer for the case of $[0,1] \subset \mathbb{R}^1$.
Let $g(x) = x - f(x)$, we want to show that $\exists x \in [0,1]$ such that $g(x) = 0$. WLOG, suppose that $g(0) = 0 - f(0) < 0$ and $g(1) = 1 - f(1) > 0$. Since x and $f(x)$ is continuous, by the *Intermediate Value Theorem*, $\exists c \in [0,1]$ such that $g(0) < g(c) = 0 < g(1)$. This means that $f(c) = c$ and hence c is a fixed point of f. □

Definition: Take $X \subseteq \mathbb{R}^n$, $f : X \to X$ is a **Contraction** if $\exists \lambda \in (0,1) \subseteq \mathbb{R}_{++}$ such that

$$\forall x, y \in X,\; \|f(x) - f(y)\| \leq \lambda \|x - y\|$$

Theorem 4.15: Banach's Fixed Point Theorem (Contraction Mapping)

Take $X \subseteq \mathbb{R}^n$ to be a closed subset of $\mathbb{R}^n$. If $f : X \to X$ is a contraction, then f has a *unique* fixed point in X[a].

[a]See Efe Ok page 176 for full proof.

Theorem 4.16: Kakutani's Fixed Point Theorem

Take $X \subseteq \mathbb{R}^n$ to be compact and convex. Let $g : X \rightrightarrows X$ to be convex-valued, closed-valued, and upper-hemi-continuous. Then the correspondence g has a fixed point in X.

Note that you won't learn Theorems 4.17 and 4.18 until Macro II.

Theorem 4.17: Contraction Mapping Theorem

Let (X, d_X) be a *complete metric space* and $f : X \to X$ is a contraction, then

(i) $\exists x^* \in X$ such that $f(x^*) = x^*$

(ii) $(x_k)_{k=1}^{\infty} \in X$, $x_{k+1} = f(x_k) \ \Rightarrow \ (x_k)_{k=1}^{\infty} \to x^*$

Theorem 4.18: Blackwell's Theorem

Let $X \subset \mathbb{R}^n$ and $C(X)$ be the space of bounded functions $f : X \to \mathbb{R}$ with the sup-metric. Let $\varphi : C(X) \to C(X)$ be a self-mapping on this space. Then if,

(i) (**Monotonicity**) $\forall x \in X,\ f(x) \leq g(x) \ \Rightarrow \ \forall x \in X,\ B(f(x)) \leq B(g(x))$

(ii) (**Discounting**) $\exists \beta \in (0,1), \forall f \in C(X)$ and $a \geq 0$ such that

$$B(f(x) + a) \leq B(f(x)) + \beta a$$

Then B is a contraction with modulus β

Remark: Notice that the common theme here is that a self-mapping on a compact and convex set yields nice properties. Based on the assumptions required, we have different strength of results. When using these for economic applications, make sure to make note of the assumptions available to you and use the results accordingly. Also note that these theorems, with the exception of Banach, only provides existence but not uniqueness. Later on (in Macro II), you will learn about the *Blackwell Conditions* for *Contraction Mapping Theorems* for more complicated cases.

5 More on Binary Relations

Recall that a binary relation is a relation B on a set X. It is commonly denoted as a subset of the Cartesian product of the set is on. ($B \subseteq X \times X$, $x, y \in X, xBy$)

Definition: A binary relation is ________ if ________

Reflexive	xBx
Irreflexive	$\neg(xBx)$
Symmetric	$xBy \ \Rightarrow \ yBx$
Asymmetric	$xBy \ \Rightarrow \ \neg(yBx)$
Anti-Symmetric	$x \neq y \wedge xBy \ \Rightarrow \ \neg(yBx)$
Total	$x \neq y \ \Rightarrow \ xBy \vee yBx$
Complete	$xBy \vee yBx$
Transitive	$xBy \wedge yBz \ \Rightarrow \ xBz$
Negatively Transitive	$\neg(xBy) \wedge \neg(yBz) \ \Rightarrow \ \neg(xBz)$
Acyclic	$\forall x_0, x_1, \ldots, x_n \in X,\ x_0Bx_1 \wedge x_1Bx_2 \wedge \cdots \wedge x_{n-1}Bx_n \ \Rightarrow \ x_0 \neq x_n$
Negatively Acyclic	$\forall x_0, x_1, \ldots, x_n \in X,\ \neg(x_0Bx_1) \wedge \neg(x_1Bx_2) \wedge \cdots \wedge \neg(x_{n-1}Bx_n) \ \Rightarrow \ x_0 \neq x_n$

Properties of Binary Relations

i. B is **asymmetric** if and only if it is **anti-symmetric** and **irreflexive**

ii. B is **complete** if and only if **total** and **reflexive**

iii. B is **acyclic** implies B is **irreflexive** and **anti-symmetric**

iv. B is **irreflexive** and **transitive** implies that B is **anti-symmetric** and **acyclic**

v. B is **total** and **transitive** implies B is **negatively transitive**

Proposition: Let P be an asymmetric binary relation on X. (a) If P is negatively transitive, then P is transitive. (b) If P is transitive, then P is acyclic.

Proof:

(a) Assume that P is negatively transitive. Then take $x, y, z \in X$ such that xPy, yPz, we want to show that xPz (so it is transitive).
Since P is negatively transitive, we know that $\neg(xPz) \wedge \neg(zPy) \ \Rightarrow \ \neg(xPy)$ is true. So the contrapositive of the statement $xPy \ \Rightarrow \ zPy \vee xPz$ is also true. Notice that P is asymmetric and yPz, so zPy must not be true, and hence xPz must be true and P is transitive.

(b) Assume that P is transitive, we want to show that $\forall x_0, x_1, \ldots, x_n \in X$, $x_0 B x_1 \wedge x_1 B x_2 \wedge \cdots \wedge x_{n-1} B x_n \Rightarrow x_0 \neq x_n$. Since P is transitive, we know that $x_0 B x_1 \wedge x_1 B x_2 \wedge \cdots \wedge x_{n-1} B x_n$ implies $x_0 B x_n$. Since P is assymetric, $x_0 B x_n \Rightarrow \neg(x_n B x_0)$ so we know that $x_0 \neq x_n$ and hence P is acyclic.

$\square$

Proposition: Let R be a complete binary relation on X.

- If R is transitive, then R is negatively transitive.
- If R is negatively transitive, then R is acyclic.

Definition: Generally, we have the following **Classes** of orders:

Weak Order	is	complete and transitive
Strict Order	is	asymmetric and negatively transitive
Weak Linear Order	is	complete, transitive, and anti-symmetric
Strict Linear Order	is	asymmetric, negatively transitive, and total

Proposition: Let R be a binary relation on X. Define $P \equiv \{(x, y) \in X \times X \mid \neg(yRx)\}$ then:

(i) R is complete if and only if P is asymmetric

(ii) R is transitive if and only if P is negatively transitive

(iii) R is negatively transitive if and only if P is transitive

(iv) R is anti-symmetric if and only if P is total

(v) Let $I \equiv \{(x, y) \in X \times X \mid xRy \wedge yRx\}$, then P is negatively transitive if and only if P is transitive AND I is transitive

Partial Orders in $\mathbb{R}^n$: Take $x, y \in \mathbb{R}^n$, then we define the following orders:

(i) $x >> y \iff \forall i \in \{1, \ldots, n\}$, $x_i > y_i$

(ii) $x > y \iff \forall i \in \{1, \ldots, n\}$, $x_i \geq y_i \wedge x \neq y$

(iii) $x \geq y \iff \forall i \in \{1, \ldots, n\}$, $x_i \geq y_i$

6 Preference Relations

Definition: A preference relation $\succsim_i$ over X is a *complete* and *transitive* binary relation on X where "$x \succsim_i y$ means "Agent i prefers x to y".

Definition: A strict preference relation is $\succ_i \equiv \{(x, y) \in X \times X \mid x \succ_i y \;\Rightarrow\; \neg(y \succ_i x)\}$
Remark: Since $\succsim$ is complete, $\succ$ is asymmetric. Since $\succsim$ is transitive, $\succ$ is negatively transitive.

Definition: An indifference relation is defined as $\sim_i \equiv \{(x, y) \in X \times X \mid x \succsim_i y \wedge y \succsim_i x\} = \{(x, y) \in X \times X \mid \neg(x \succ_i y) \wedge \neg(y \succ_i x)\}$

Definition: A **Cobb-Douglas** preference on $\mathbb{R}^2_+$ is a relation
$\succsim_i \equiv \{(x, y) \in \mathbb{R}^2_+ \times \mathbb{R}^2_+ \mid x_1x_2 \geq y_1y_2\}$.

Definition: An **Additively separable, linear preference** is a binary relation on $X \equiv \mathbb{R}^N_+$ defined by:

$$\succsim_i \equiv \left\{(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^N_+ \mid \sum_{i=1}^{N} \lambda x_k \geq \sum_{i=1}^{n} \lambda y_k,\ \lambda \in \mathbb{R}_{++}\right\}$$

Definition (Equivalence Relation): An **Equivalence Relation** is a binary relation that is *reflexive*, *transitive*, and *symmetric*.

Definition (indifference class): The **Indifference Class** of $x \in X$ for agent i is defined as $I_i(x) \equiv \{y \in X \mid x \sim_i y\}$. If $x \succ_i y$, then $\forall a \in I_i(x), b \in I_i(y)$, we have $a \succ_i b$.

Proposition[14]**:** $\forall x, y \in X$ such that $x \sim_i y$, $I_i(x) \equiv I_i(y)$

Definition (Maximal Elements): Let B be a binary relation on X. Let $Y \subseteq X$. We say that $x \in Y$ is an **Maximal Element** of (a binary relation) on Y if

$$\forall y \in Y,\ yBx \;\Rightarrow\; xBy$$

Further, we denote the set of maximal elements of Y for the relation B as $M(Y, B)$.

[14] Take $z \in I_i(x)$, we have $z \sim_i x$. We also know that $x \sim_i y$, so $z \sim_i y$, so $z \in I_i(y)$.

Proposition: $\forall S \subseteq X$, $\forall x \in S$ the following are equivalent:

(i) x is maximal for $\succ_i$ in S

(ii) x is maximal for $\succsim_i$ in X

(iii) $\forall y \in S$, $\neg(y \succ_i x)$

(iv) $\forall y \in S$, $x \succsim_i y$

Proposition: Let B be an asymmetric binary relation on the set X. $M(S,B)$ is non-empty for any finite, non-empty subset $S \subseteq X$ **if and only if** B is acyclic

Proof:

"$\Leftarrow$":
Assume that B is acyclic. Suppose otherwise that $M(S,B)$ is empty. Since S is non-empty, take $y \in S$. Since $M(S,B) = \varnothing$, we know that $\exists y_2 \in S$ such that $y_2 B y$. Since $M(S,B) = \varnothing$, we know $y_2 \notin M(S,B)$ so $\exists y_3 \in S$ such that $y_3 B y_2$. Since S is finite and non-empty, suppose that $|S| = n \in \mathbb{N}$. then since $M(S,B) = \varnothing$, $\forall k \in \{1,\dots,n\}, \exists y_k \in S$ such that $y_k B y_{k-1}$. Since $|S| = n$, $y_n B y$. But B is acyclic, so, by contradiction, $M(S,B) \neq \varnothing$.

" $\Rightarrow$ ":
Assume that $M(S,B) \neq \varnothing$, $|S| = n \in \mathbb{N}$. Suppose otherwise that B is cyclic so that $\forall i \in \{1,\dots,n\}, y_i \in S$, $y_i B y_{i+1} \Rightarrow y_1 = y_n$. Since $M(S,B) \neq \varnothing, \exists k \in \{1,\dots,n\}$ such that $y_k B y_{k-1}$ and $\forall y \in S$, $y_k B y$. But if B is cyclic, we have $y_{k+1} B y_k$, so by contradiction, B is acyclic.

$\square$

Corollary: The set of maximals of *strict* relations on a *finite* set is non-empty.

Definition: Take the metric space (X, d_X) and an asymmetric binary relation B on X. We say that B is **Continuous** on X if $\forall x, y \in X$ such that xBy, $\exists \varepsilon \in \mathbb{R}_{++}$ such that $\forall z \in N_\varepsilon^{d_X}(x), \forall w \in N_\varepsilon^{d_X}(y)$, we have zBw.

Proposition: Take the *compact* metric space (X, d_X). If B is acyclic and continuous, then $M(X,B)$ is non-empty.

Remark: Intuitively, a "rational" agent would choose the *maximal* element out of their choice set based on their *preference relation*.

Definition: A **Choice rule** C over the set X is a mapping $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ or $C : 2^X \setminus \varnothing \rightrightarrows X \setminus \varnothing$ such that $\forall\, non-empty\, S \subseteq X$, $C(S) \subseteq S$.

Definition: Take the set X, a binary relation B on X, and a choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$. We say that B **Rationalizes** C if $\forall S \subseteq X$, $S \neq \varnothing$, $C(S) = M(S, B)$. We say that C is **Rationalizable** if such B exists.

Proposition: A choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ is rationalizable if and only if it is rationalizable by an acyclic relation.

Proof:

" $\Rightarrow$ ":
If C is rationalizable by an assymetric relation B, then B is acyclic since there is at least one maximal element. If C is only rationalizable by a non-asymmetric relation B, we can construct $P \equiv \{(x, y) \in X \times X \mid \neg(yBx)\}$. By construction, P is asymmetric. Notice that $M(X, B) = M(X, P)$ since $\forall x \in M(X, B), \forall y \in Y$, $yBx \;\Rightarrow\; xBy$. The contrapositive of this statement is $\neg(xBy) \;\Rightarrow\; \neg(yBx)$ which is equivalent to $xPy \;\Rightarrow\; yPx$, so $x \in M(X, P)$, and hence P rationalizes C and is acyclic.

"$\Leftarrow$": This direction is immediate.

$\square$

Remark: Take $X = \{x, y, z\}, C(\{y, z\}) = \{y\}$, $C(\{x, y\}) = \{x\}$, $C(\{x, y, z\}) = y$, we want to show that no acyclic relation P can rationalize C.
$M(\{y, z\}, P) = \{y\} \;\Rightarrow\; z \notin M(\{y, z\}, P)$, so $yPz \wedge \neg(zPy)$
$M(\{x, y\}, P) = \{x\} \;\Rightarrow\; y \notin M(\{x, y\}, P)$ so $xPy \wedge \neg(yPx)$
This means that the relationship is actually cyclic.

Definition: A choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ is said to satisfy **Condition $\boldsymbol{\alpha}$** if $\forall S \subseteq T \subseteq X$, $S \neq \varnothing$,

$$x \in C(T) \cap S \;\Rightarrow\; x \in C(S)$$

Example: If the best athlete in the world is an American, then the best American athlete is that same person. Meaning, what you choose in a larger set will be your choice in the smaller subset.

Definition: A choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ is said to satisfy **Condition $\boldsymbol{\gamma}$** if $\forall S, T \subseteq X$, $S, T \neq \varnothing$

$$x \in C(S) \wedge x \in C(T) \;\Rightarrow\; x \in C(S \cup T)$$

Proposition: A choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ on a finite set X is *rationalizable* if and only if C satisfies conditions α and γ.

Proof: $\alpha \wedge \gamma \;\Rightarrow\;$ rationalizability (John Duggan Notes)

Define the base relation $R_c \equiv \{(x, y) \in X \times X \mid x \in C(\{x, y\})\}$. Notice that R_c is reflexive and complete (because choices are non-empty)

Lemma (ASB 1.1.1): A choice rule C is rationalizable if and only if it is rationalizable by the base weak preference relation R_c.

Proof:

Assume that C is rationalizable, we want to show that it is rationalizable by R_c. Recall that if C is rationalizable, it is rationalizable by an acyclic relation R.
Define a complete relation $R_R \equiv \{(x, y) \in X \times X \mid xRy \text{ or } \neg(yRx)\}$ where R is an acyclic relation. Notice that R_R is a complete relation.

(i) We want to show that if C is rationalizable by R, then it rationalizable by R_R
$\forall Y \subseteq X$, $C(Y) = M(Y, R)$. We know that $x \in M(Y, R) \iff \forall y \in Y$, xR_Ry[a]
Since R_R is complete, we know that $x \in M(Y, R_R)$. This means that $M(Y, R) = M(Y, R_R)$, so $C(Y) = M(Y, R) \;\Rightarrow\; C(Y) = M(Y, R_R)$, so C is rationalizable by R_R

(ii) We want to show that $R_R = R_c$
$\forall x, y \in X$, $xR_Ry \iff x \in M(\{x, y\}, R) \iff x \in C(\{x, y\}) \iff xR_cy$[b] $\square$

Back to our main proof

"$\Leftarrow$": Assume that α and γ hold, we want to show that C is rationalizable by R_c.

Take $Y \subseteq X, x \in C(Y), y \in Y$. Then $x \in C(\{x, y\})$[c]. By the previous lemma, we then know that xR_cy, meaning that $x \in M(Y, R_c)$.

We have now shown that $C(Y) \subseteq M(Y, R_c)$. Now we need to show that $M(Y, R_c) \subseteq C(Y)$, and this is where we require γ to be true.

Take $x \in M(Y, R_c)$, since X is a finite set, so is Y. Let $|Y| = n \in \mathbb{N}$, we can order $Y = \{y_1, y_2, \ldots, y_n\}$. Then $\forall i \in \{1, \ldots, n\}, x \in C(\{x, y_i\})$, so $x \in \bigcap_{i=1}^{m} C(\{x, y_i\})$. Since γ is assumed true, $x \in C(\bigcup_{i=1}^{m}\{x, y_i\}) = C(Y)$.
So we know that C is rationalizable by R_c.

" $\Rightarrow$ ": Suppose that C is rationalizable by R_c, we want to show that α and γ hold.

Take $Y \subseteq Z \subseteq X$, take $x \in Y\, s.t.\, x \in C(Z)$, we want to show that $x \in C(Y)$. Since $x \in C(Z)$, $x \in M(Z, R_c) \cap Y$. Since $Y \subseteq Z$, $x \in M(Y, R_c)$ so α holds.

Take $Y, Z \subseteq X$, take $x \in C(Y) \cap C(Z)$, so $x \in M(Y, R_c)$ and $x \in M(Z, R_c)$. So we have that $\forall y \in Y,\ z \in Z,\ (xR_cy \vee \neg(yR_cx) \wedge (xR_cz \vee \neg(zR_cx))$, so $x \in M(Y \cup Z, R_c)$ and $x \in C(Y \cup Z)$, so we know that γ holds.

$\square$

[a]Definition of maximal element. $x \in M(Y, R) \Rightarrow \forall y \in Y,\ yRx \Rightarrow xRy$, so either $\neg(yRx)$ or xRy is always true.
[b]This last part is just the definition of R_c
[c]Because α is assumed to be true.

Definition[15]**:** A choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ satisfies **Condition $\boldsymbol{\beta}$** if $\forall S \subseteq T \in 2^X \setminus \varnothing$

$$x \in C(S) \wedge y \in C(S) \wedge x \in C(T) \ \Rightarrow\ y \in C(T)$$

Proposition: A choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ on a finite set X is rationalizable by the standard preference relations ($\succsim$ and $\succ$) if and only if C satisfies both α and β.

[15]Idea: if something is elevated, so are all things linked to it.

Proof: $\alpha \wedge \beta \Rightarrow$ rationalizability

"$\Leftarrow$": Assume that C satisfies α and β, we want to show that C is rationalizable by $\succsim$ and $\succ$.

Lemma: $\alpha \wedge \beta \Rightarrow \gamma$
Proof:
Let $S, T \in 2^X \setminus \varnothing$ and take $x \in C(S) \cap C(T)$, we want to show that $C(S \cup T)$. Let $Z \equiv S \cup T$. Suppose, WLOG, that $S \cap C(Z) \neq \varnothing$, then take $y \in S \cap C(Z)$. By α, since $S \subset Z$, we know that $y \in C(S)$.

By β, since $x \in C(S) \wedge y \in C(S) \wedge y \in C(Z)$, we know that $x \in C(Z)$. So $x \in C(S) \wedge x \in C(T) \Rightarrow x \in C(Z \equiv S \cup T)$, so γ holds.
□

Back to our main proof

By lemma, we know that $\alpha \wedge \gamma \iff C$ is rationalizable by the complete relation R_c defined earlier.

Claim: $\alpha \wedge \beta \Rightarrow R_c$ is transitive
Take $x, y, z \in X$ such that $xR_c y \wedge yR_c z$, we want to show that $xR_c z$

(i) $y \in C(\{x, y, z\})$, then[a] $y \in C(\{x, y\})$. By assumption[b], $x \in C(\{x, y\})$. So $C(\{x, y\}) = \{x, y\}$. By $\beta, x \in C(\{x, y, z\})$, so by α, $x \in C(\{x, z\})$ and $xR_c z$

(ii) Similarly, $z \in C(\{x, y, z\})$, then $z \in C(\{y, z\})$. By assumption $y \in C(\{y, z\})$. By β, $y \in C(\{x, y, z\})$, so $xR_c z$.

(iii) Suppose that $y, z \notin C(\{x, y, z\})$, then $C(\{x, y, z\}) = \{x\}$. By α, $x \in C(\{x, z\})$, so $xR_c z$

"$\Leftarrow$": Assume that C is rationalizable by $\succsim$ and $\succ$, we want to show that β holds
Take $S, T \in 2^X \setminus \varnothing$, $S \subseteq T$, take $x \in C(S), y \in C(S), x \in C(T)$, we want to show that $y \in C(T)$. Since C is rationalizble by $\succsim$, $C(T) = M(T, \succsim)$, $C(S) = M(S, \succsim)$ and $y \in M(S, \succsim)$, since $x \in M(S, \succsim)$, we know $x \succsim y$ and $y \succsim x$. Since $x \in M(T, \succsim), \forall t \in T, t \succsim x \Rightarrow x \succsim t$. We know that $y \succsim x$ so $\forall t \in T$, $t \succsim y \Rightarrow t \succsim x \Rightarrow x \succsim t \Rightarrow y \succsim t$, so $y \in M(T, \succsim)$. Hence β holds. □

[a] Because α is assumed to be true.
[b] $xR_c y$

Definition (WARP): A choice rule $C : 2^X \setminus \varnothing \to 2^X \setminus \varnothing$ satisfies the **Weak Axiom of Revealed Preferences** if $\forall S, T \in 2^X \setminus \varnothing$, we have:

$$x \in C(S) \wedge y \in S \wedge y \notin C(S) \wedge y \in C(T) \;\Rightarrow\; x \notin T$$

Proposition: WARP $\iff \alpha \wedge \beta$

Proof: WARP $\iff \alpha \wedge \beta$

" $\Rightarrow$ ":

Assume that α, β holds but $WARP$ is violated. Take $x \in C(S)$, $y \in S$, $y \notin C(S)$, $y \in C(T)$, $x \in T$. Then by α, we know that $x \in C(S) \wedge y \in C(T) \;\Rightarrow\; x, y \in C(S \cap T)$, then by β, we know that $x \in C(S \cap T) \wedge x \in C(S) \;\Rightarrow\; y \in C(S)$. But We had assumed that $y \notin C(S)$. Hence, by contradiction, $\alpha \wedge \beta \;\Rightarrow\; WARP$.

"$\Leftarrow$":

Assume that $WARP$ is satisfied, but β is not. Then $\exists S, T \subseteq X$ such that $S \subset T \subset X$. Take $x \in C(S)$, $y \in C(S)$, $x \in C(T)$, since β is violated, $y \notin C(T)$. Then since $x \in C(T)$, $y \in T$, $y \notin C(T)$, $y \in C(S)$, by $WARP$, $x \notin C(S)$, but x was assumed to be in $C(S)$. Hence, by contradiction, $WARP$ implies β.

Assume that $WARP$ is satisfied, but α is not. Then $\exists S \subseteq T \subseteq X$ and $x \in C(T)$ such that $x \notin C(S)$, $x \in S$, $y \in C(S)$, meaning that $y \in S$, $y \in T$. Since $y \in C(S)$, $x \in S$, $x \notin C(S), x \in C(T)$, by $WARP$, we know that $y \notin T$, but we already assumed that $y \in T$. Hence, by contradiction, $WARP$ also implies α.

□

Definition: A function $U : X \to \mathbb{R}$ represents $\succsim$ $(\succ)$ over X if $\forall x, y \in X$

$$x \succsim (\succ) y \iff U(x) \geq (>) U(y)$$

Proposition: If U represents $\succsim$, then for any strictly increasing function $f : \mathbb{R} \to \mathbb{R}, V(x) \equiv f(U(x))$ also represents $\succsim$ as well (utility function is ordinal[16]).

[16] An ordinal utility function specifically means that the utility representations denote the "order", but the difference in differences means nothing. For example, an increase in utility from 5 to 10 cannot be compared to increase in utility from 10 to 15. Later on, we learn about quasi-linearity where utility seems less ordinal because of the dense-ness of the real numbers.

Proof: Utility Functions are Identical up to Str. Inc. Transformations

Since U represents $\succsim$, $\forall x, y \in X$, $U(x) \geq U(y) \iff x \succsim y$. Since f is a strictly increasing function, $x \succsim y \iff U(x) \geq U(y) \iff f(U(x)) \geq f(U(y)) \iff V(x) \geq V(y)$. So $V(x)$ represents $\succsim$ as well. $\square$

6.1 Sufficient Conditions for Utility Representation

6.1.1 X is Finite

Definition: Take X, $\succsim$, we say that $x \in X$ is minimal in X if $\forall y \in X$, $y \succsim X$. In any finite X, there is a minimal element.

Proof: Any Finite Set X has a Minimal Element

If $|X| = 1$, $x \in X$ is minimal since $\succsim$ is complete.
Suppose $|X| = n \in \mathbb{N}$ has a minimal element $x_k \in X$, take $x_{n+1} = X_n \cup \{x_{n+1}\}$ where $x_{n+1} \notin X_n$. Since $\succsim$ is complete, we know that $x_k \succsim x_{n+1} \vee x_{n+1} \succsim x_k$. If the former is true, than by transitivity, x_{n+1} is a minimal element. If the latter is true, then x_k is still the minimal element. $\square$

Claim: Take $\succsim$ on a finite set X. Then $\succsim$ has utility representation $U : X \to \mathbb{N}$

Proof: Utility Representation when X is Finite

Let X_1 be the indifference class of elements that are minimal. Define $U(x) = 1$, $\forall x \in X_1$. By the previous lemma, $X_1 \neq \varnothing$. Now suppose we have $X_1, \ldots, X_k$ such that $U(x) = i$, $\forall x \in X_i$. Let X_{k+1} be the minimal indifference class in $X \setminus \bigcup_{i=1}^{k} X_i$. If $X \setminus \bigcup_{i=1}^{k} X_i \neq \varnothing$, then by lemma, $X_{k+1} \neq \varnothing$. Since X is finite, $\exists j \in \mathbb{N}$ such that $X \setminus \bigcup_{i=1}^{j} X_i = \varnothing$

" $\Rightarrow$ ": Suppose $x \succsim y$, then $\forall y \in X_k$, $x \in X \setminus \bigcup_{i=1}^{k-1} X_i$, so $U(y) = k \leq U(x)$

"$\Leftarrow$": Suppose $U(x) \geq U(y)$, then $x \in X \setminus \bigcup_{i=1}^{U(y)-1} X_i$, so $x \succsim y$

$\square$

6.1.2 X is Countably Infinite

(Intuition) Since X is countable, it can be ordered, so the previous proof should follow.

Proposition: Let X be a countably infinite set, then $\succsim$ on X can be represented by some function $U : X \to (-1, 1)$

Proof: $\succsim$ on Countably Infinite Set X is Representable

Let $X = \{(x_n)_{n=1}^{\infty}\}$ and let $U(x_1) = 0$. Suppose that this sequence is ordered such that $U(x_i) \geq U(x_j) \iff x_i \succsim x_j$.

If $m \in \{1, \ldots, k\}$ such that $x_{k+1} \sim x_m$, we have $U(x_{k+1}) = U(x_m)$.

If not, by transitivity of $\succsim$, $\exists l, h \in \{1, \ldots, k\}$ such that $\{U(x_l) \mid x_{k+1} \succ x_l\}$ is "strictly below" $\{U(x_h) \mid x_h \succ x_{k+1}\}$.

So $\max\left\{\{U(x_l) \mid x_{k+1} \succ x_l\} \cup \{-1\}\right\} < \min\left\{\{U(x_h) \mid x_h \succ x_{k+1}\} \cup \{1\}\right\}$.

Since R is complete, we have

$$\max\left\{\{U(x_l) \mid x_{k+1} \succ x_l\} \cup \{-1\}\right\} < U(x_{k+1}) < \min\left\{\{U(x_h) \mid x_h \succ x_{k+1}\} \cup \{1\}\right\}$$

$\square$

Definition: A **Lexicographical Preference** $\succsim_L$ is a preference relation that satisfies a ordered-coordinate wise preference relation. In other words, take $x, y \in X^2$,

$$x \succsim_L y \iff x_1 \succ y_1 \vee (x_1 \succsim y_1 \wedge x_2 \succsim y_2)$$

Proposition: Lexicographical preferences on $[0, 1]^2$ does not have a utility representation.

Proof: Lexicographical Preference on $[0, 1]^2$ is NOT Representable

Suppose otherwise that there exists a utility function $U : [0, 1]^2 \to \mathbb{R}$ that represents $\succsim_L$, then $\forall a \in [0, 1]$, $U(a, 1) > U(a, 0)$.

Define a function q such that $q(a) \in \mathbb{Q} \cap (U(a,0), U(a,1))$.

Notice that $q(a)$ must be a strictly increasing function in a on $[0,1]$ so $q(a)$ must be an injection from $[0,1]$ to $\mathbb{Q}' \subset \mathbb{Q}$.

But $[0,1]$ is *uncountably* infinite whereas $\mathbb{Q}'$ is at most *countably* infinite, so such $q(a)$ must not exist. Hence, by contradiction, U does not exist. $\square$

6.1.3 Separable and Convex X with continuous $\succsim$

Definition: Take $(X, d_X, \succsim)$, a subset $S \subseteq X$ is **Order-Dense** according to $\succsim$ if $\forall x, y \in X$ such that $x \succ y$, $\exists z \in S$ such that $x \succ z \succ y$

Theorem 6.1: Birkhoff 1

Take $(X, d_X, \succsim)$. If X contains a countable, order-dense subset, then $\succsim$ is representable.

The proof of this theorem can be found in Efe. Ok. P.104

Theorem 6.2: Birkhoff Condition

Take $(X, d_X, \succsim)$, then $\succsim$ on X is representable if and only if X contains a countable Y such that $\forall x, y \in X \setminus Y$ such that $x \succ y$, $\exists z \in Y$ such that $x \succ z \succ y$

Definition: Take $(X, d_X, \succ)$, we say that $\succ$ on X is **Continuous** if $\forall x, y \in X$ *s.t.* $x \succ y$, $\exists \varepsilon \in \mathbb{R}_{++}$ such that

$$\forall w \in N_\varepsilon^{d_X}(x),\ z \in N_\varepsilon^{d_X}(y),\ w \succ z$$

Definition *(Sequential Characterization)*: $\succsim$ is **Continuous** on X if $\forall (x_n, y_n)_{n=1}^\infty \in X$ such that $(x_n, y_n)_{n=1}^\infty \to (x, y)$ we have

$$\forall n \in \mathbb{N},\ x_n \succsim y_n \ \Rightarrow \ x \succsim y$$

Proposition: The two definitions are equivalent[17].

Proof: The two definitions are equivalent

" $\Rightarrow$ ":
Assume that $\forall x, y \in X$ such that $x \succsim y$, $\exists \varepsilon \in \mathbb{R}_{++}$ such that $\forall w \in N_\varepsilon^{d_X}(x)$, $z \in N_\varepsilon^{d_X}(y)$, $w \succsim z$. Take $(x_n, y_n)_{n=1}^\infty \in X$ such that $\forall n \in \mathbb{N}$, $x_n \succsim y_n$ and $(x_n, y_n)_{n=1}^\infty \to (x, y)$.

Suppose otherwise that $y \succ x$, then by assumption, $\exists \varepsilon \in \mathbb{R}_{++}$ such that $\forall w \in N_\varepsilon^{d_X}(x)$, $z \in N_\varepsilon^{d_X}(y), z \succ w$, but since $(x_n, y_n)_{n=1}^\infty \to (x, y)$, $\forall \varepsilon_x \in \mathbb{R}_{++}$, $\exists N \in \mathbb{N}$ such that $\forall n > N$, $(x_n, y_n) N_{\varepsilon_x}^{d_X}(x, y)$. Now take $\varepsilon^* = \min\{\varepsilon, \varepsilon_x\}$, then $y_n \succ x_n$, but by construction, $x_n \succsim y_n$, so by contradiction, we have $x \succsim y$

"$\Leftarrow$":
Assume $\forall (x_n, y_n)_{n=1}^\infty$ such that $(x_n, y_n) \to (x, y)$ and $\forall n \in \mathbb{N}, x_n \succsim y_n$, then $x \succsim y$. We want to show that $\exists \varepsilon \in \mathbb{R}_{++}$ such that $\forall w \in N_\varepsilon^{d_X}(x)$, $z \in N_\varepsilon^{d_X}(y)$, $w \succ z$.

Suppose otherwise that $z \succsim w$, then take a convergent sequence $\forall w_n \in N_\varepsilon^{d_X}(x)$, $z_n \in N_\varepsilon^{d_X}(y)$ such that $(w_n, z_n) \to (x, y)$. Then by assumption, $y \succsim x$. But by assumption, $x \succsim y$, so by contradiction, we must have that $w \succ z$ $\square$

Theorem 6.3: Debreu 1954

Let (X, d_X) be a separable and convex space and $\succsim$ a continuous preference relation on X. Then there exists a continuous function $U : X \to \mathbb{R}$ that represents $\succsim$.

Proof 6.3: Debreu 1954

Lemma: Take a continuous weak preference $\succsim$ on a convex set $x \subseteq \mathbb{R}^n$, then $\forall x \succ y$, $\exists z \in X$ such that $x \succ z \succ y$.

Proof:
Suppose otherwise. Let I be the interval connecting x and y. By construction, $I \subseteq X$.

[17] And later on, we can use this to show that if a preference is represented by a continuous utility function, then the preference must be continuous. The converse is not true.

Let us construct a sequence of "midpoints" in the following way:

Step 1: Take the "midpoint" of x and y: $m_0 = \frac{x+y}{2}$

Step 2: If $x \succ m_0 \succ y$, then the proof is done. Otherwise, define the next midpoint as:

$$m_1 = \begin{cases} \frac{y+m_0}{2} & , m_0 \succ x \\ \frac{x+m_0}{2} & , m_0 \prec y \end{cases}$$

Step 3: We will iteratively repeat this process until $\exists k \in \mathbb{N}$ such that $x \succ m_k \succ y$ where

$$m_k = \begin{cases} \frac{y+m_{k-1}}{2} & , m_{k-1} \succ x \\ \frac{x+m_{k-1}}{2} & , m_{k-1} \prec y \end{cases}$$

Since $\mathbb{R}$ is complete, we know such m_k exists.

Back to our main proof

Proposition: Take a convex subset of $\mathbb{R}^n$, $X \subset \mathbb{R}^n$ and a continuous weak preference relation $\succsim$ on X. Then $\succsim$ has a utility representation.

Proof:
Since X is separable, we can take a countable and dense subset $Y \subseteq X$. We know that $\exists u : Y \to (-1, 1)$ that represents $\succsim$ on Y since Y is countable (proven here). Then $\forall x \in X$, define $V(x) \equiv \sup\{v(z) \mid z \in Y, x \succ z\}$ where $V(x) = -1$ if $\{z \in Y \mid x \succ z\} = \varnothing$. So if $x \succ y$, we want to show that $V(x) > V(y)$.

By the lemma we just proved, $\exists z \in X$ such that $x \succ z \succ y$. By continuity of $\succsim$, we know that $\exists \varepsilon \in \mathbb{R}_{++}$ such that $\forall w \in N_\varepsilon^{dx}(z)$, $x \succ w \succ y$. Since Y is dense in X, $\exists z_1 \in N_\varepsilon^{dx}(z) \cap Y$ such that $x \succ z_1 \succ y$. In fact, $\forall i \in \mathbb{N}$, $\exists z_i \in N_\varepsilon^{dx}(z) \cap Y$ such that $x \succ z_1 \succ z_2 \succ \cdots \succ y$. So $V(x) \geq v(z_1) > v(z_2) \geq V(y)$.

The indifference part is automatic. Since $\succsim$ is continuous on X. If we take 2 sequences that approach the same point, the limit of their utility representation will equal and they will be indifferent. □

7 Choice under Uncertainty

In general, there are 3 established approaches of studying choice under uncertainty. We will study them in the following order:

1. Expected Utility Theory (aka Von Neuman & Morgansten 1944)

2. Subjective Utility Theory (Savage 1954)

3. Horserace/Roulette Theory (Anscombe-Aumann 1953)

7.1 A Brief Mention of Measure Theory

Naturally, we can think about uncertainty as a probability function that assigns elements of a set X to the interval $[0, 1]$. This is generally denoted as $p : X \to [0, 1]$ such that $\sum_{x \in X} p(x) = 1$. However, one would quickly realize that this only works for countable sets. One attempt, commonly referred to as the **simple probability approach**, at solving this problem in the case of uncountably infinite sets is to reduce it down to a countable/finite subset and then assign the countable events with probabilities.
This might seem enough in day-to-day use, but if you wish to study advanced econometrics or game theory, you will need the measure theoretic approach.

Definition: An **Algebra** on X is a collection of subsets of X that is closed under finite union, intersections, and complements.

Definition: A **σ-algebra** on X is a collection of subsets of X that is closed under countable union, intersections, and complements.

Definition: The **Borel σ-algebra** on X is the smallest $\sigma-$algebra that contains all open subsets of X.

Definition: A probability measure is a *size function* that *measures* the size of the elements of $B(X)$, the Borel algebra of X. Let $\mu : B(X) \to [0, 1]$ and $A, B \subseteq X$, we have

(i) $\mu(\varnothing) = 0$

(ii) $\mu(A) = 1 - \mu(X \setminus A)$

(iii) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$

7.2 Expected Utility Theory

Take a finite set of states $X = \{x_1, x_2, \ldots, x_n\}$ and define a probability measure $P : X \to [0, 1]$ such that $\sum_{x \in X} p(x) = 1$. Define the set of **lotteries** $L(X)$ as the set of vectors where each "spot" is a probability assigned to an element in X. Let $\delta_{x_i} \in L(X)$ denote the **degenerate lottery** where x_i happens with probability 1. Assume that x_i's are ordered such that $\delta_{x_1} \succsim \delta_{x_2} \succsim \cdots \succsim \delta_{x_n}$

Rational Choice Axiom: Monotonicity
We say that a preference $\succsim$ satisfies *monotonicity* if

$$a\delta_{x_1} + (1-a)\delta_{x_n} \succsim b\delta_{x_1} + (1-b)\delta_{x_n} \iff a \geq b$$

Rational Choice Axiom: Continuity
We say that a preference $\succsim$ satisfies *Continuity* if $\forall p \in L(X)$, $\exists a \in [0, 1]$ such that

$$a\delta_{x_1} + (1-a)\delta_{x_n} \sim p$$

Note that the Expected Utility Theorem assumes the following 2 axioms:

Rational Choice Axiom: Independence
$\forall p, q, r \in L(X)$ and $\forall \alpha \in (0, 1)$

$$p \succsim q \iff \alpha p + (1-\alpha)r \succsim \alpha q + (1-\alpha)r$$

Rational Choice Axiom: Archimedian Property
Given $p, q, r \in L(X)$ *s.t.* $p \succ q \succ r$, $\exists \alpha \in (0, 1)$ such that $\alpha p + (1-\alpha)r \sim q$.

Claim (MIT final 2005): If $\succsim$ satisfies *monotonicity* and *continuity*, then $\succsim$ satisfies the *Archimedian property*. The proof is simple and left to the reader.

Since X is finite, we have a utility representation $u : [p, r] \to [0, 1]$ such that $U(x) = \alpha$ if $x \notin \{p, r\}$ and $U(r) = 0, U(p) = 1$. Since such function is continuous and over a finite set[18] by the Weierstrass Theorem of Maximum, we know the maximizer and minimizer of u within X exist.

[18]Notice that finite sets are closed and bounded in $\mathbb{R}^n$

Theorem 7.1: Expected Utility Theorem

Take a finite set X, the weak preference relation $\succsim$ on $L(X)$ satisfies the *Archimedian Property* and *Independence* **if and only if** $\exists\, u : X \to \mathbb{R}$ such that $\forall p, q \in L(X)$

$$p \succsim q \iff \underbrace{\sum_{x \in X} p(x)u(x)}_{\text{Expected value of } U \text{ according to lottery } p} \geq \sum_{x \in X} q(x)u(x)$$

Moreover, the utility representation u is unique up to positive affine transformations[a].

[a]See proof in MWG page 176 and Rubenstein page 96.

Extension: But what if X is not finite? We can try to use the **simple lottery** approach and define $L_S(X) \equiv \{p \in [0,1]^X \mid \|\{x \in X \mid p(x) > 0\}\| \in \mathbb{N}, \sum_{x \in X} p(x) = 1\}$. Then the theorem applies on $L_S(X)$.

7.3 Risk Attitude on an Ordered Set

Motivating Example: The Allais (1953) Paradox

Consider 4 lotteries: p_1, p_2, q_1, q_2 with the following payoff structure:

	p_1	p_2	q_1	q_2
0	0%	1%	89%	90%
1000	100%	89%	11%	0%
5000	0%	10%	0%	10%

Empirically, people prefer p_1 to p_2 but q_2 to q_1. If *independence* holds, we should have been able to reduce the problem to

	p_1	p_2	q_1	q_2
0		1%		1%
1000	11%		11%	
5000		10%		10%

So the comparison between the p's and q's should have been the same. So what went wrong?

Take $X \subseteq \mathbb{R}$ where X has a distribution $F_X : X \to [0,1]$ such that

$$F_X(x) = \begin{cases} \sum\limits_{y \in (-\infty, x]} p(y) & \text{if } X \text{ is a finite set} \\ \mu\left((-\infty, x]\right) & \text{if } X \text{ is uncountably infinite and } \mu \text{ is a measure over } B(X) \end{cases}$$

We will focus most of our discussions on the discrete case

Take $\succsim_L$ (note that this is not the Lexicographical preference previously discussed as an aside.) represented by a utility function u on X so that $u(F) \geq u(F') \iff F \geq F'$ where $F, F' \in L(X)$ are lotteries and $U(F) = \sum\limits_{x \in X} p(x)u(x)$ or $U(F) = \int_X f(x)u(x)dx$ where $F_X(x) = \int_x f(t)dt$

Definition: The **Expected Value** (EV) of F over a closed interval $[a,b]$ is

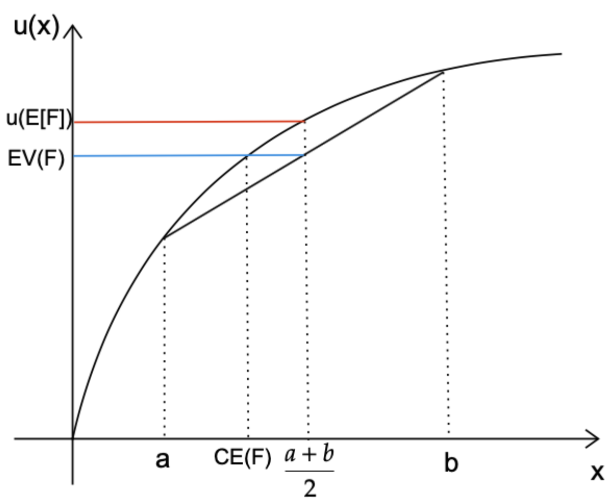
$$U(F) = \begin{cases} \sum\limits_{x \in [a,b]} p(x)u(x) & \text{if } X \text{ is finite} \\ \int\limits_a^b f(x)u(x)dx & \text{if } X \text{ if uncountably infinite} \end{cases}$$

We say that Agent i is <u>Risk Neutral</u> if $\forall F \in L([a,b])$, we have $\underline{E[U(F)] = U(E[F])}$
Risk Averse $E[U(F)] \leq U(E[F])$
Risk Loving $E[U(F)] \geq U(E[F])$

so a risk-averse (loving) utility function is concave (convex).

Definition: The **Certainty Equivalent** of a lottery for agent i, denoted as $CE(F, u_i) \in \mathbb{R}$ is $x \in \mathbb{R}$ such that $u_i(x) = u_i(CE[F]) = EV(F, u_i)$.

Definition: The **Risk Premium**[19] is $RP_i = u_i^{-1}(EV(F, u_i)) - CE(F, u_i)$. In the following example, the risk premium is $\frac{a+b}{2} - CE(F, u_i)$. So for a risk-averse agent, $RP_i > 0$.



[19]How much you are willing to pay to avoid the lottery.

But is there a way for us to figure out whether someone is **more** risk-averse than others? Arrow-Pratt answers this question for us.

Definition: Take a C^2 utility function $u : x \to \mathbb{R}$. The **Arrow-Pratt Coefficient of Absolute Risk Aversion (CRRA)** is defined as

$$r \equiv -\frac{u''(x)}{u'(x)}$$

Notice that u is strictly monotonic so $u'' = 0 \iff u(x) = c \in \mathbb{R}$. Also note that this is a rate and so is scale-irrelevant. We also define a *constant absolute risk aversion (CARA)* utility function as $u(x) = -e^{-cx}$.

Definition: Take a C^2 utility function $u : x \to \mathbb{R}$. The **Arrow-Pratt Coefficient of Relative Risk Aversion** is defined as

$$r^R \equiv -\boldsymbol{x}\frac{u''(x)}{u'(x)}$$

Note that **RRA** is scale-relevant. Also note that CRRA utility functions are used in Macroeconomics quite a bit as $\dfrac{c^{1-\sigma}-1}{1-\sigma}$.

Definition: Take $[a, b] \subset \mathbb{R}_+$, let u_i, u_j be strictly increasing and C^2 on $[a, b]$. We say that agent i is **More Risk-Averse** than j if any of the 4 following (equivalent) conditions is true:

(i) $\forall x \in [a, b]$, $r_i(x) \geq r_j(x)$

(ii) There exists a *weakly concave* function $g : \mathbb{R} \to \mathbb{R}$ such that $u_i(x) = g(u_j(x))$

(iii) $\forall F \in L([a, b])$, $CE(F, u_i) \leq CE(F, u_j)$

(iv) For any non-degenerate (not sure outcome) lottery F and degenerate outcome x, if agent i prefers F to x, then j prefers lottery F to x

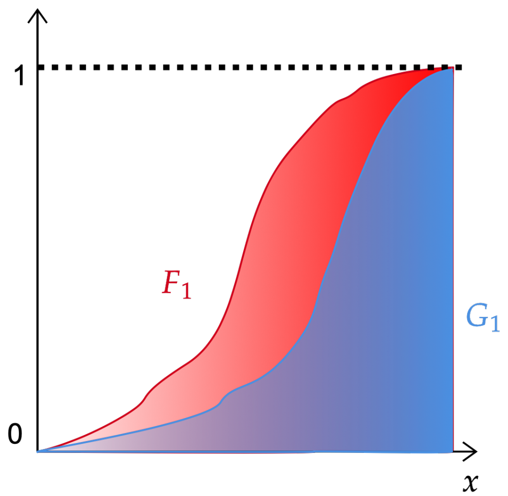
7.3.1 Stochastic Dominance

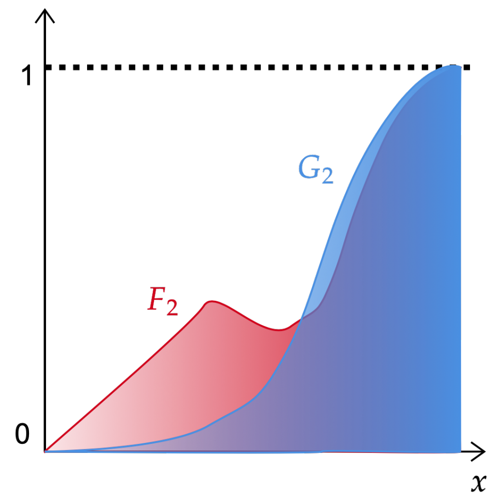
Definition (FOSD): Take $F, G \in L([a, b])$, $F \neq G$, where $F(x)$ and $G(x)$ are the CDF of the lottery ordered by the preference relation. We say that F **First-Order Stochastically Dominates** G if $\forall x \in [a, b] \subset \mathbb{R}$, $F(x) \leq G(x)$ with strict inequality at some $x \in [a, b]$.

Definition (SOSD): Take $F, G \in L([a,b])$, we say that F **Second-Order Stochastically Dominates** G if ALL risk-averse agent would prefer F over G.

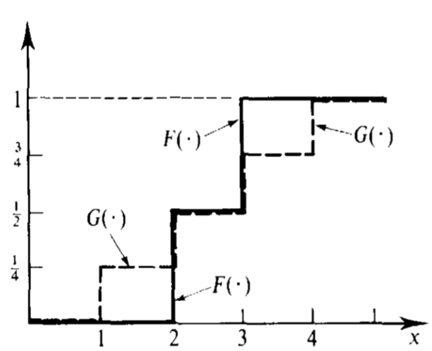
Alternative Definition (SOSD): Take $F, G \in L([a,b])$, we say that F **Second-Order Stochastically Dominates** G if $\forall \boldsymbol{x} \in [a,b]$, $\int\limits_{-\infty}^{x} F(t)\,dt \leq \int\limits_{-\infty}^{x} G(t)\,dt$.

Consider the following two sets of comparisons. In the graph on the left, G_1 $FOSD$ F_1, so intuitively, G_1 $SOSD$ F_1 (see area to the left of the curves). In the graph on the right, G_2 does NOT $FOSD$ F_2, but G_2 $SOSD$ F_2.





Definition (MPS): A **Mean-Preserving Spread** is a *further* randomization within a lottery that has mean 0. This means that the expected potential outcomes are the same but there is more risk added[20].
Here is a graphical example of *Mean-Preserving Spread* from MWG p.198:



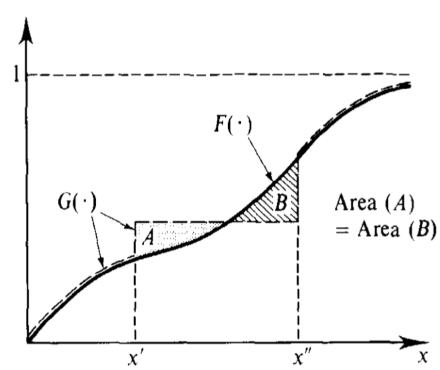


Figure 6.D.3 (left)
$G(\cdot)$ is a mean-preserving spread of $F(\cdot)$.

Figure 6.D.4 (right)
$G(\cdot)$ is an elementary increase in risk from $F(\cdot)$.

[20]Idea: Choose a mean and randomize out from there, so we have a "re-randomization" in both directions of the first distribution. Variance does not change.

Proposition: Let F, G be lotteries <u>with the same mean</u> then

$$F\ SOSD\ G \iff G \text{ is a mean-preserving spread of } F$$

7.4 Subjected Utility Theory (Savage 1954)

Motivating Example: Take the *state* of a random berry in the park to be the set $S = \{Poisonous, Non-poisonous\}$, the *outcomes* of the **interaction** with said berry to be $Z = \{happy, hungry, sick\}$, and the *actions*[21] $A = \{eat, not-eat\}$, where eat and not eat are functions that map the states of the berry to set of outcomes.

$eat(poisonous) = sick,\ eat(non-poisonous) = happy,\ not-eat(s \in S) = hungry.$
Savage proposes that when we observe revealed preferences, we are choosing within A based on our **subjective** beliefs of the probability space $\Pi(S)$.

Formally, take S as the set of states, Z the set of outcomes, and A the set of actions that are $f : S \to Z$. We want to study $\succsim$ on A. For $f \in A$ that are degenerate, define $\succsim_Z$ such that $x \succsim_Z y \iff \forall x, y \in A,\ x \succsim y$.

Definition: $\forall B \subseteq S$, define $\succsim_B$ as the **Conditional preference relation** over the set of actions, restricted to states in B. then $\forall f, g \in A$

$$f \succsim_B g \iff f' \succsim g' \ s.t.\ f|_B = f'|_B \wedge g|_B = g'|_B \wedge f'|_{B^c} = g'|_{B^c}$$

Definition: $B \subseteq S$ is **Null** if $\forall f, g \in A,\ f \sim_B g$

Notation: $\forall f \in A,\ B \subseteq S,\ f|_B$ is $g : B \to Z$ defined by $\forall b \in B,\ g(b) = f(b)$

Axiom (SEU1): $\forall f, f', g, g' \in A$, $\forall B \subseteq S$ and $f|_B = f'|_B \wedge g|_B = g'|_B \wedge f|_{B^c} = g|_{B^c} \wedge f'|_{B^c} = g'|_{B^c}$, then

$$f \succsim g \iff f' \succsim g'$$

Axiom (SEU2): $\forall$ non-null $B \subseteq S$, $\forall f, g \in A$, $\forall b \in B\ s.t.\ f(b) = x,\ g(b) = y$

$$f \succsim_B g \iff x \succsim y$$

[21] Actions are functions that map states to outcomes.

So preferences of one thing over the other is not state-dependent

Axiom (SEU3): Take $x, y, x', y' \in Z$, $f, g, f', g' \in A$, $B, C \subseteq S$ such that

(i) $x \succ y$, $x' \succ y'$

(ii) $f|_B = x$, $f|_{B^c} = y$, $f'|_B = x'$, $f'|_{B^c} = y'$

(iii) $g|_C = x$, $g|_{C^c} = y$, $g'|_C = x'$, $g'|_{C^c} = y'$

then

$$f \succsim g \iff f' \succsim g'$$

Axiom (Sure-thing Principle) (SEU4): Take $B \subseteq S$, then $\forall b \in B$

$$\begin{aligned} f \succ_B g(b) &\Rightarrow f \succsim_B g \\ g(b) \succ_B f &\Rightarrow g \succsim_B f \end{aligned}$$

Axiom (SEU5): Take $f, g, f', g' \in A$ such that $f \succ g$ and $\forall x \in Z$, then there exists a partition $\Pi(S)$ of S such that $\forall B \subseteq \Pi(S)$ we have

(i) $f'|_B = x$, $f'|_{B^c} = f|_{B^c} \Rightarrow f' \succ g$

(ii) $g'|_B = x$, $g'|_{B^c} = g|_{B^c} \Rightarrow f \succ g'$

Theorem 7.2: Subjective Expected Utility Theorem (Savage 1954)

If $\succsim$ on A satisfies the axioms SEU 1-5, then there exists a **unique** probability measure Π and density function Π' defined over the $\sigma-$algebra generated by S AND a **bounded** function $u : Z \to \mathbb{R}$ (unique up to affine transformations) such that $\forall f, g \in A$, $\forall B, C \subseteq S$, and $\forall x, y \in Z \; s.t. \; f|_B = g|_C = x$, $f|_C = g|_B = y$ such that $x \succ y$, the following are true

$$f \succsim g \iff \int_{s \in S} \Pi'(s) u(f(s)) ds \geq \int_{s \in S} \Pi'(s) u(g(s)) ds$$

and

$$\Pi(B) > \Pi(C) \iff f \succ g$$

7.5 Horserace-Roulette (Anscombe-Aumann)

Hopefully, you have realized that the von-Neuman-Morganstern approach has run into plenty of walls (paradoxes) in its lifetime, and that the subjective expected utility approach is so versatile that it is hardly tractable. Luckily, Anscombe-Aumann proposed a middle ground. They suggested that the probability space of certain events is actually comprised of 2 parts: subjective and objective. The subjective part follows the Savage approach and allows us free reign on agents' beliefs. The objective part follows the vNM approach where certain events have known probabilities.

Formally, we define the actions as probability functions so that $f \in A$, $f : S \to P(Z)$. Intuitively, this means that agents choose actions based on what they believe the resulting objective probability space would be.

Theorem 7.3: Horserace/Roulette Expected Utility Theorem

The preference relation $\succsim$ over the set of actions A is representable by an utility function $U : S \to \Pi(s)$ if and only if the following axioms hold:

(A1) $\forall f, g \in A,\ f \succsim g \vee g \succsim f$

(A2) $\forall f, g, h \in A,\ f \succsim g \wedge g \succsim h \ \Rightarrow\ f \succsim h$

(A3) $\forall f, g, h \in A\ s.t.\ f \succ g \succ h,\ \forall a, b \in [0,1],\ (af + (1-a)h \succ g) \wedge (g \succ bf + (1-b)h$

(A4) $\forall f, g, h \in A\ s.t.\ f \succ g,\ \forall a \in [0,1],\ af + (1-a)h \succ ag + (1-a)h$

The utility representation in this framework is defined as:

$$
\begin{aligned}
f_Z :& S \to P(Z) \\
\Pi :& P(Z) \to [0,1]\ s.t. \sum_{s \in S} \Pi(s) = 1 \\
u :& Z \to \mathbb{R}
\end{aligned}
$$

such that

$$U(f) = \sum_{p \in P(Z)} \left(\Pi(p) \cdot \left(\sum_{s \in S} f_Z(s) u\left(z(s)\right) \right) \right)$$

8 Consumer Theory

Having built a foundation for preference representations, we shall now study how agents make decisions based on their preferences. Take a finite set of m goods M ($|M| = m \in \mathbb{N}$) and some good $k \in M$. Our agents have the set of choices $X \subseteq \mathbb{R}^m_+$ where $x \in X$ is called a **consumption bundle**. For simplicity, we will consider $X = \mathbb{R}^m_+$ and define the weak preference relation $\succsim_i$ for agent i on X. Consumption of good k is denoted by $x_k \in \mathbb{R}_+$.

Properties a preference relation $\succsim$ can have on X:

1. **Monotonicity** (Recall the definition of partial orders in $\mathbb{R}^n$):

   (i) A preference relation $\succsim$ on X is **non-decreasing** at $y \in X$ if

   $$\forall x \in X, x \geq y \;\Rightarrow\; x \succsim y$$

   (ii) A preference relation $\succsim$ on X is **monotonic** at $y \in X$ if

   $$\forall x \in X, x >> y \;\Rightarrow\; x \succ y$$

   (iii) A preference relation $\succsim$ on X is **strongly monotonic** at $y \in X$ if

   $$\forall x \in X, x \geq y \wedge x \neq y \;\Rightarrow\; x \succ y$$

   A preference relation is said to be **locally non-satiated** at $y \in X$ if

   $$\forall \varepsilon \in \mathbb{R}_{++}, \exists x \in N^{d_X}_{\varepsilon}(y) \; s.t. \; x \succ y$$

2. **Continuity**:
   A preference relation $\succsim$ on X is said to be **continuous** if $\forall x, y \in X$, $\exists \varepsilon \in \mathbb{R}_{++}$ such that $\forall w \in N^{d_X}_{\varepsilon}(x)$ and $\forall z \in N^{d_X}_{\varepsilon}(y)$

   $$x \succ y \iff w \succ z$$

   Alternatively, we can define $\succsim$ to be continuous if $\forall x \in X$, its *upper-contour* and *lower-contour* sets are closed. i.e.,$\forall x \in X$, $\succsim_H(x) \equiv \{y \in X \mid y \succsim x\}$ and $\succsim_L(x) \equiv \{y \in X \mid y \precsim x\}$ are closed. Additionally, if $\forall x \in X$, $\succsim_H(x)$ ($\succsim_L(x)$) is closed, we say $\succsim$ is upper(lower)-semi-continuous.

3. **Convexity**:
   A preference relation $\succsim$ on X is said to be (**strictly**) **convex** if $\forall x, y \in X$ and $\forall \lambda \in [0,1]$

$$y \succsim x \iff \lambda y + (1-\lambda)x (\succ) \succsim x$$

   Notice that this is equivalent to saying that $\forall x \in X$, $\succsim_H(x)$ is convex

Definition: A preference relation $\succsim$ on $\mathbb{R}^m_+$ is said to be **Homothetic** $\forall \alpha \in \mathbb{R}_{++}$

$$x \succsim y \;\Rightarrow\; \alpha x \succsim \alpha y$$

Definition: A preference relation $\succsim$ on $\mathbb{R}^m_+$ is said to be **Quasi-Linear** in the **Numeraire** (good 1 in M) if $\forall \alpha \in \mathbb{R}$, $\forall \beta \in \mathbb{R}_{++}$

$$\begin{aligned}
&x \succsim y \;\Rightarrow\; x + \alpha(1,0,\ldots,0) \succsim y + \alpha(1,0,\ldots,0) \\
&\text{AND} \\
&x + \beta(1,0,\ldots,0) \succ x
\end{aligned}$$

Now that the preferences are well-defined, let's define the rest of the agent's problem. Let $p \in \mathbb{R}^m_{++}$[22] be an exogenous price vector and let $w \in \mathbb{R}_{++}$ denote the wealth of the agent. We can then define the budget set as:

$$B(p,w) \equiv \{x \in \mathbb{R}^m_+ \mid px^\top \leq w\}$$

Notice that the budget set, by construction, is convex, and we can thus define our standard problem of utility maximization in statics as

$$\max_{x \in B(p,w)} u(x)$$

where we are guaranteed a solution by the Weierstrass Theorem of Maximum.

Definition: The **Walrasian/Individual Demand Correspondence** is hence defined as $x : \mathbb{R}^m_{++} \times \mathbb{R}_+ \rightrightarrows \mathbb{R}^m_+$ such that

$$x(p,w) \equiv \underset{x \in B(p,w)}{argmax}\; u(x)$$

[22] In Eguia's notations, he always let p be a $1 \times m$ vector and x be a $m \times 1$ vector. To practice matrix algebra, I will try to let x be a $1 \times m$ vector just like other vectors. If I miss something, please let me know.

Proposition (assumptions): We will generally assume $X = \mathbb{R}^m_+$ and $\succsim$ is *continuous* and *locally non-satiated* on X.

Properties a Demand Correspondence Can Have:
Take a continuous, locally non-satiated preference relation $\succsim$ on $X = \mathbb{R}^m_+$. The individual demand correspondence $x : \mathbb{R}^m_{++} \times \mathbb{R} \rightrightarrows \mathbb{R}^m_+$ satisfies the following properties:

(i) The demand correspondence is **homogeneous**[23] **of degree** 0 in (p, w). i.e.,

$$\forall \alpha \in \mathbb{R}_{++},\, x(\alpha p, \alpha w) = x(p, w)$$

(ii) (Walras' law)[24] The budget set is exhausted. i.e., $px^\top(p, w) = w$

(iii) $x(p, w)$ is upper-hemi-continuous (by Berge's Theorem of Maximum)

(iv) If $\succsim$ is a convex preference relation, then $x(p, w)$ is convex-valued

(v) If $\succsim$ is a *strictly* convex preference relation, then $x(p, w)$ is a continuous function[25].

Definition: Suppose that $u : X \to \mathbb{R}$ is locally differentiable at $x \in \mathbb{R}^m_{++}$, we define the **Marginal Rate of Substitution** of goods i and j to be

$$MRS_{ij} \equiv \frac{\partial u(x)/\partial x_i}{\partial u(x)/\partial x_j}$$

For an interior solution ($x \in \mathbb{R}^N_{++}$), it can be shown that $MRS_{ij} = \frac{p_i}{p_j}$.

Definition: The Lagrange Multiplier λ is the **Shadow Price** of the corresponding constraint.

Definition: Take a price vector $\bar{p}$, the maximizing bundle under price $\bar{p}$ is defined as

$$x_{\bar{p}}(w) \equiv x(p, w)|_{p=\bar{p}}$$

Definition: The image of $x_{\bar{p}}(w)$ is called the **Wealth Expansion Path** under price $\bar{p}$.

[23] Homogeneous of degree 1 means $x(\alpha p, \alpha w) = \alpha x(p, w)$

[24] This is a result of local non-satiation

[25] Since $x(p, w)$ is uhc, once $\succsim$ was restricted to strictly convex, the graph of each point is a singleton, and hence $x(p, w)$ is a continuous function.

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $\forall k \in M$, the **Wealth/Income Effect** for the k^{th} good is

$$\frac{\partial x_k(p, w)}{\partial w}$$

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $\forall k \in M$, a good k is **Normal** at (p, w) if its income effect at (p, w) is *non-negative.*

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $\forall k \in M$, a good k is **Inferior** at (p, w) if its income effect at (p, w) is *strictly negative.*

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $\forall k \in M$, the *Marshallian* **price effect** of p_l on the k^{th} good is

$$\frac{\partial x_k(p, w)}{\partial p_l}$$

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $\forall k \in M$, a good k is **Giffen** at (p, w) the price effect of p_k on good k at (p, w) is positive[26].

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $\forall k \in M$, the **Price Elasticity**[27] of good k with respect to the price of good l (p_l) is:

$$\varepsilon_{kl}(p, w) \equiv \frac{\partial x_k(p, w)}{\partial p_l} \cdot \frac{p_l}{x_k(p, w)}$$

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $\forall k \in M$, the **Income Elasticity/Elasticity of Demand** of good k with respect to w is:

$$\varepsilon_{kw}(p, w) \equiv \frac{\partial x_k(p, w)}{\partial w} \cdot \frac{w}{x_k(p, w)}$$

Claim: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, if x is a correspondence that is homogeneous of degree 0 with respect to (p, w), then

$$\forall k \in M,\ \varepsilon_{kw} + \sum_{l \in M} \varepsilon_{kl} = 0$$

Intuitively, this means that the sum of price effects of all other goods has to be equivalent, in quantity, to the wealth effect.

[26]Giffen goods are a theoretical subset of inferior goods. See The Duality Paradox for a detailed discussion.
[27]If $l = k$, we call it the own-price elasticity. If not, we call it the cross-price elasticity

8.1 Marshallian Demand

Definition: A **Consumer's Choice Rule** is $C : \mathbb{R}^M_{++} \times \mathbb{R}_+ \to X$ such that $\forall (p, q) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $C(p, w) \in B(p, w)$.

Definition: $\succsim$ on X rationalizes the **Consumer Choice Function** $C(p, w)$ if $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $C(p, w) = M\left(B(p, w), \succsim\right)$.

We will further simplify by saying *Consumer C* for agents whose choice functions are rationalized by $\succsim$.

Definition: For consumer C, we define the **Revealed Preference** as:

$$\succsim_R \equiv \{x \succsim y \mid \exists (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+ \; s.t. \; x, y \in B(p, w) \; \wedge \; x = C(p, w)\}$$

Definition: Let consumer C be rationalized by $\succsim$. We say that $\succsim$ satisfies **Consumer-WARP** (C-WARP) if $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$,

$$x = C(p, w) \wedge py' \leq w \wedge y = C(\tilde{p}, \tilde{w}) \;\Rightarrow\; \tilde{p}x' > \tilde{w}$$

Remark: Readers should realize that what sets *C-WARP* apart from our original definition of WARP is that this version requires the assumption of *local non-satiation*. In this version, if X is chosen over Y under (p, w) but Y is chosen under (p', w'), it must be that X is not affordable under (p', w').

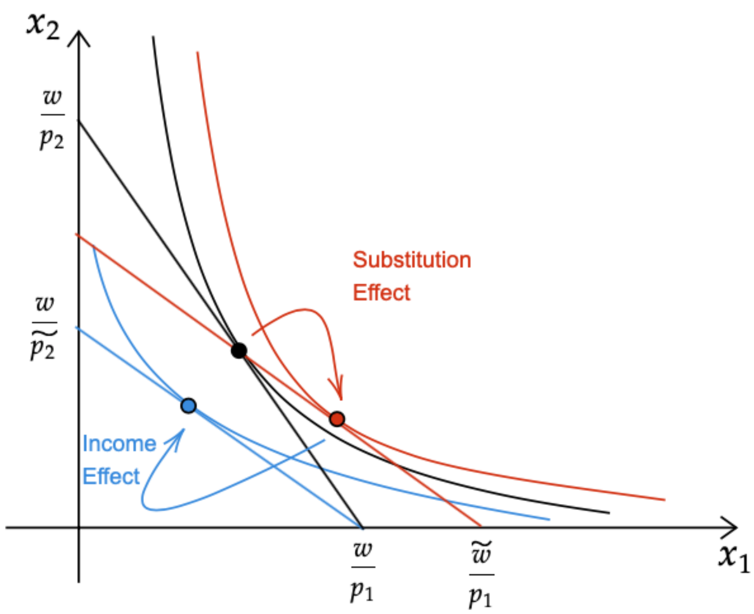
This is more restrictive as it implies that an increase in w' can potentially bring X back into the picture. Given fixed p', an increase in w' necessarily means more purchasing power and the ability to consume more of something. In other words, C-WARP requires the consumer to prefer more of certain goods, and hence the original WARP plus local non-satiation is equivalent to C-WARP.

Definition: $\forall (p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$ and for all consumer C. $\forall k, l \in M$, the **Substitution Effect** of price p_l on the demand for good k (x_k) at (p, w) is

$$s_{kl}(p, w) \equiv \overbrace{\frac{\partial x_k(p, w)}{\partial p_l}}^{\text{Price Effect}} - \left(\overbrace{\underbrace{-x_l(p, w)}_{=\frac{\partial w}{\partial p_l}} \cdot \underbrace{\frac{\partial x_k(p, w)}{\partial w}}_{\text{change in } x_k \text{ per change in } w}}^{\text{Wealth/Income Effect}}\right)$$

The following graph dissects price effects (black dot to blue dot) into substitution effects (black dot to red dot) and income effects (red dot to blue dot) for the changes in optimal consumption from (p, w) to $(\tilde{p}, w)$ where p_2 increased and p_1 is unchanged:

Substitution: Given the new price ratio and a new income $\tilde{\omega}$[28], the consumer substitutes to a new bundle at a higher indifference curve (black to red).
Income: Given the original income, the consumes has to decrease the consumption as compared to the consumption at $(\tilde{p}, \tilde{\omega})$ (red to blue).
Price: These two effects combined (black to blue).



Definition: If we further assume that $C : \mathbb{R}^m_{++} \times \mathbb{R}_+ \to X$ is twice differentiable at (p, w), we can define the **Slutsky Substitution Matrix** as:

$$S(p, w) \equiv \begin{pmatrix} s_{1,1}(pw) & \cdots & s_{1m}(p, w) \\ \vdots & \ddots & \vdots \\ s_{m,1}(pw) & \cdots & s_{mm}(p, w) \end{pmatrix}$$

and we know that $pS(p, w) = 0_{1\times m}$.

Proposition: Let $\succsim$ on X be locally non-satiated and continuous. Assume that $x(p, w)$ is differentiable. Then

$$\sum_{k=1}^{m} p_k s_{kl} = 0 = \sum_{l=1}^{m} p_l s_{kl}(p, w)$$

Proposition: If $m = 2$, then the Slutsky's substitution matrix $S(p, w)$ is *symmetric*. If C satisfies C-WARP, then $S(p, w)$ is *negative semi-definite.*

[28]Such that the original bundle (black) is affordable

Proposition: A consumer choice function C is rationalizable by $\succsim_R$ if and only if $\forall(p, w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, $S(p, w)$ is symmetric and negative semi-definite.

Definition (SARP): A consumer choice function C satisfies the **Strong Axiom of Revealed Preferences** if for any list of price and wealth $\{(p^1, w^1), (p^2, w^2), \ldots, (p^\tau, w^\tau)\} \in \mathbb{R}^m_{++} \times \mathbb{R}_+, \tau \in \mathbb{N}, h \in \{1, \ldots, \tau-1\}$ such that $C(p^{h+1}, w^{h+1}) \neq C(p^h, w^h)$ AND $p \cdot C(p^{h+1}, w^{h+1}) \leq w^h$, then

$$\sum_{k=1} p^\tau C_k(p^1, w^1) > w^\tau$$

Remark: According to Chat-GPT (which cites Big Kreps) and Robin Danko, the difference between C-WARP and SARP is that SARP requires transitivity while C-WARP does not.

Definition (GARP): A consumer choice function C satisfies the **General Axiom of Revealed Preferences** if for *every* ordered subset of the natural numbers $\{i, j, k, \ldots, r\} \subset \mathbb{N}$ such that

$$\begin{gathered} p_i \cdot x_j \leq p_i \cdot x_i \\ p_j \cdot x_k \leq p_j \cdot x_j \\ \ldots \\ p_r \cdot x_i \leq p_r \cdot x_r \end{gathered}$$

then it must be true that all these inequalities are actually **equalities**.

Theorem 8.1: Afriat's Theorem

If a consumer choice function satisfies GARP, then it is rationalizable by a strongly monotonic, continuous, and convex $\succsim$ (and can be represented by some piece-wise linear, continuous, strictly monotonic, and concave utility function).

Recall that our typical consumer problem is

$$\max_{x \in \mathbb{R}^m_+} u(x) \ s.t. \ px^\top = w$$

with the typical solution

$$x(p, w) \equiv \underset{x \in \mathbb{R}^m_+}{argmax}\, u(x) \ s.t. \ px^\top = w$$

Now, by Berge's Theorem of Maximum, we know that $\max u(x(p,w))$ is continuous and $x(p,w)$ is upper-hemi-continuous. If we go with the open set definition of uhc, it is not hard to see that $\max u(x(p,w))$ is continuous in (p,w). **This is huge**, because this means we don't actually need to solve for $x(p,w)$ every time just to understand how utilities change.

Definition (Indirect Utility Function $v(p,w)$): Given that the preference is continuous and locally non-satiated. The **Indirect Utility Function**, derived from the preference's utility representation, is the function $v : \mathbb{R}^m_{++} \times \mathbb{R}_+ \to \mathbb{R}$ defined as $v(p,w) = \max u(x(p,w))$.

Proposition: The indirect utility function $v(p,w)$ has the following properties:

(i) $v(p,w)$ is homogeneous of degree 0 in $(p,w) \in \mathbb{R}^m_{++}]RR_+$

(ii) $v(p,w)$ is strictly increasing in w and non-increasing in p_k[29]. $\forall k \in \{1,\ldots,m\}$

(iii) $v(p,w)$ is quasi-convex. $\forall \bar{v} \in \mathbb{R}$, $\{(p,w) \mid v(p,w) \leq \bar{v}\}$ is convex

(iv) $v(p,w)$ is continuous[30] in p and w

Important Proposition (Roy's Identity): Let a consumer choice function be rationalized by a *locally non-satiated* and *strictly convex* preference $\succsim$. Suppose that $\succsim$ has the utility representation $u(x)$ and the indirect utility function $v(p,w)$. If $v(p,w)$ is locally differentiable at $(\hat{p},\hat{w}) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$, then $\forall k \in M$

$$x_k(p,w) = -\frac{\partial v(\hat{p},\hat{w})/\partial p_k}{\partial v(\hat{p},\hat{w})/\partial w}$$

To see this, notice that the indirect utility function is the composite of the utility function and the demand function (since the preference is assumed to be strictly convex). Suppose that $v(p,w)$ is differentiable at $(p,w) \in \mathbb{R}^m_{++} \times \mathbb{R}_+$. By the envelope theorem, we know that

$$\frac{\partial v(p,w)}{\partial p_k} = \frac{\partial u(x(p,w))}{\partial p_k} = -\lambda(x_k) \tag{1}$$

where λ is the shadow price of relaxing the budget constraint by an arbitrarily small amount, so $\lambda = \frac{\partial u(x(p,w))}{\partial w} = \frac{\partial v(p,w)}{\partial w}$. So we can rewrite equation (1) and "solve" for $x_k(p,w)$ as:

$$\frac{\partial v(p,w)}{\partial p_k} = -\frac{\partial v(p,w)}{\partial w} x_k \;\Rightarrow\; x_k(p,w) = -\frac{\partial v(p,w)/\partial p_k}{\partial v(p,w)/\partial w}$$

[29]This is because the maximizer could be constant in x_k

[30]By Berge's Theorem of Maximum

8.2 Hicksian Demand and Duality

Throughout this semester, we have always referred to the *typical consumer problem* where agents maximize their utility subject to a constraint. Like what Savage (1954) did to the expected utility theory with his idea of subjective expected utility theory, we will look at the flip side of the coin - the **duality framework**.

The Hicksian dual problem is about changing the utility maximization problem into an expenditure minimization problem. It posits that consumers can make their decisions by choosing a bundle that minimizes their expenditure, given that the bundle satisfies a certain level of utility.

Formally, the Hicksian problem is:

$$\min_{x \in \mathbb{R}^m_+} px^\top \; s.t. \; u(x) \geq \bar{u}$$

The solution $h(p, \bar{u}) \in \mathbb{R}^m_+$ to this problem is called the **Hicksian/Compensated Demand**.

Definition: The **Expenditure Function** $e : \mathbb{R}^m_{++} \times \mathbb{R} \to \mathbb{R}_+$ is a function of the price vector $p \in \mathbb{R}^m_{++}$ and utility $\bar{u}$ such that $e(p, \bar{u}) = ph^\top(p, \bar{u})$.

Proposition: Suppose $u(\cdot)$ is a continuous utility function representing a *locally non-satiated* preference relation $\succsim$, then the expenditure function has the following properties:

(i) $e(p, u)$ is homogeneous of degree 1 in p

(ii) $e(p, u)$ is strictly increasing in u, and non-decreasing in p_l

(iii) $e(p, u)$ is concave in p

(iv) $e(p, u)$ is continuous in p and u

Proposition: Suppose $u(\cdot)$ is a continuous utility function representing a *locally non-satiated* preference relation $\succsim$. Let $h(p, u)$ denote the Hicksian demand correspondence, then we have

1. $h(p, \bar{u}) = h(p, u(x(p, w))) = h(p, v(p, w)) = x(p, w)$
   $e(p, u(x, w)) = w$

2. $x(p, p \cdot h(p, \bar{u})) = x(p, e(p, \bar{u})) = x(p, w) = h(p, \bar{u})$
   $v(p, e(p, \bar{u})) = \bar{u}$

Proposition: Suppose $u(\cdot)$ is a continuous utility function representing a *locally non-satiated* preference relation $\succsim$, then the Hicksian demand correspondence has the following properties:

(i) $h(p,u)$ is homogeneous of degree 0 in p

(ii) Delivers no excess utility. $u(h(p,\bar{u})) = \bar{u}$

(iii) $h(p,u)$ is upper-hemi-continuous in p and u

(iv) If $\succsim$ is strictly convex, then $h(p,w)$ is a function

Proposition: Suppose $u(\cdot)$ is a continuous utility function representing a *locally non-satiated* and strictly convex preference relation $\succsim$, then the Hicksian demand function for good k is, $\forall k \in M$, $(p,\bar{u}) \in \mathbb{R}^m_{++} \times \mathbb{R}$

$$h_k(p,\bar{u}) = \frac{\partial e(p,\bar{u})}{\partial p_k}$$

Proof:

The Hicksian problem can potentially be solved by the Lagrangian

$$\mathcal{L}(h,\lambda) = -ph' + \lambda(u(h) - \bar{u})$$

Assuming that NDCQ is satisfied everywhere, we have the first order conditions:

$$\frac{\partial \mathcal{L}(h)}{\partial h_1} = p_1 - \lambda \frac{\partial u(h)}{\partial h_1} = 0$$
$$\vdots$$
$$\frac{\partial \mathcal{L}(h)}{\partial h_m} = p_m - \lambda \frac{\partial u(h)}{\partial h_m} = 0$$

So

$$\begin{aligned}\frac{\partial e(p,\bar{u})}{\partial p_k} &= h_k(p,\bar{u}) + \sum_{l\neq k} \underbrace{p_l}_{=\frac{\partial u(h)}{\partial h_l}} \cdot \frac{\partial h_l}{\partial p_k} \\ &= h_k(p,\bar{u}) + \underbrace{\lambda \sum_{l \neq k} \frac{\partial u(h)}{\partial h_l} \cdot \frac{\partial h_l(p,\bar{u})}{\partial p_k}}_{\text{This is 0 since the budget is exhausted}} = h_k(p,\bar{u})\end{aligned}$$

$\square$

If we further assume that h is differentiable with respect to p, then we also have

$$\frac{\partial h_k}{\partial p_l} = \frac{\partial^2 e(p,\bar{u})}{\partial p_k \partial p_l}$$

where we can think about the gradient as both a Jacobian of all the h_k and a Hessian of the expenditure function:

$$J_p(h) = \begin{pmatrix} \frac{\partial h_1}{\partial p_1} & \cdots & \frac{\partial h_1}{\partial p_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial p_1} & \cdots & \frac{\partial h_m}{\partial p_m} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 e(p,\bar{u})}{\partial p_1 \partial p_1} & \cdots & \frac{\partial^2 e(p,\bar{u})}{\partial p_1 \partial p_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 e(p,\bar{u})}{\partial p_m \partial p_1} & \cdots & \frac{\partial^2 e(p,\bar{u})}{\partial p_m \partial p_m} \end{pmatrix} = H_p(e)$$

Proposition: Let $u(\cdot)$ be a continuous utility function representing a *locally non-satiated* and strictly convex preference relation $\succsim$. If $h(p,\bar{u})$ is differentiable at $(p,\bar{u}) \in \mathbb{R}^m_{++} \times \mathbb{R}$, then $J_p(h(p,\bar{u}))$ is *negative semi-definite* and *symmetric*. Because it is equivalent to the *Slutsky Substitution matrix*, we also know that $p \cdot J_p(h(p,\bar{u}) = 0$

Definition: Assume that $h(p,\bar{u}))$ is differentiable, we say that goods k and l are **Substitutes** at $(p,\bar{u})$ if

$$\frac{\partial h_k(p,\bar{u})}{\partial p_l} \geq 0$$

Definition: Assume that $h(p,\bar{u}))$ is differentiable, we say that goods k and l are **Complements** at $(p,\bar{u})$ if

$$\frac{\partial h_k(p,\bar{u})}{\partial p_l} < 0$$

8.2.1 The Duality Paradox

Unlike the Marshallian demand correspondences, Hicksian demand correspondences do not allow the existence of Giffen goods. That is, the Hicksian demand always satisfies the law of demand so that $\forall p, p'$,

$$(p' - p) \cdot [h(p',\bar{u}) - h(p,\bar{u})] \leq 0$$

But how does that fit in the *duality* framework? Do Marshallian Giffen goods "break" the duality framework? The answer is NO, and we will see it in a little bit. The key here is that Hicks discusses/develops the framework for studying consumer preferences while intentionally omitting income effects. This was justified because Hicks took a "top-down" approach of using market demand (more on that later) to inform how individual demand behaves.

Proposition (Slutsky's Equation): Let $u(\cdot)$ be a continuous utility function representing a *locally non-satiated* and strictly convex preference relation $\succsim$. Let $h(p,u) = x(p,w)$ be differentiable. Then $\forall (p,w) \in \mathbb{R}^m_{++} \times \mathbb{R}_{++}$, we have

$$\underbrace{\frac{\partial h_k(p,v(p,w))}{\partial p_l}}_{\text{Hicksian Price Effect}} = \underbrace{\frac{\partial x_k(p,w)}{\partial p_l}}_{\text{Marshallian Price Effect}} + \underbrace{\frac{\partial x_k(p,w)}{\partial w}}_{\text{Income Effect}} \cdot x_l(p,w)$$

Now consider the case where we only have two goods l and k. Rewriting this equation and do some algebra, we get

$$\underbrace{\frac{p_l}{x_k}\overbrace{\frac{\partial x_k(p,w)}{\partial p_l}}^{\text{Marshallian Price Effect}}}_{\text{Marshallian Cross Price Elasticity}} = \underbrace{\overbrace{\frac{\partial h_k(p,v(p,w))}{\partial p_l}}^{\text{Hicksian Price Effect}} \cdot \frac{p_l}{x_k}}_{\text{Substitution Effect}} - \underbrace{\overbrace{\frac{\partial x_k(p,w)}{\partial w}}^{\text{Parts of Income Elasticity}} \cdot \overbrace{\frac{w - p_k x_k(p,w)}{x_k}}^{p_l x_l(p,w)}}_{\text{Income Effect}}$$

Rewriting this equation to get own price elasticity to gain an insight to our duality paradox.

$$\underbrace{\frac{p_k}{x_k}\overbrace{\frac{\partial x_k(p,w)}{\partial p_k}}^{\text{Marshallian Price Effect}}}_{\text{Marshallian Cross Price Elasticity}} = \underbrace{\frac{\partial h_k(p,v(p,w))}{\partial p_k} \cdot \frac{p_k}{x_k}}_{\text{Hicksian Cross Price Elasticity}} - \underbrace{\frac{\partial x_k(p,w)}{\partial w}}_{\text{Income Effect}} \cdot p_k$$

Notice that this solves our duality paradox. First, notice that the Hicksian price effect is, for a normal good (income elasticity/effect ≥ 0), more pronounced than the Marshallian price effect (Marshallian is more negative than Hicksian[31]).
Next, notice that if a good is inferior (income elasticity < 0) and the magnitude of the effect times price is large enough, than the own price elasticity of good k would be positive, meaning that it violates the law of demand (Giffen goods). But this does not affect the fact that Hicksian elasticity is negative, hence the duality paradox is a non-problem.

The Slutsky's Equation is how economists historically knew where to look for Giffen goods. An example is potato consumption in Ireland during the Irish potato famine. The historical background at the time allowed for income effects to "dominate" substitution effects and create "Giffen goods".

[31] Notice that this means the inverse demand curve, with which we commonly graph, is steeper for Hicksian than Marshallian for a normal good and vice versa for an inferior good.

8.2.2 Back to Duality

Now that our duality crisis has been solved, we can refocus on learning more about the Hicksian demand framework. We want to focus our discussion on 2 main things:

1. Recover the expenditure function from the Marshallian demand correspondence x_k. (But we will run into integrability problems. See Big Kreps 11.5)

2. Recover $\succsim$ from the expenditure function.

Recall that under the duality framework, $h(p, \bar{u}) = x(p, w)$ and

$$\frac{\partial h_k(p, \bar{u})}{\partial p_l} = \frac{\partial x_k(p, w)}{\partial p_l} = s_{kl}(p, e(p, \bar{u}))$$
$$\Rightarrow \quad J_p(h_k) = H_p(e) = S(p, w)$$

If $H_p(e)$ is negative semi-definite and symmetric, then so is $S(p, w)$. This means that $x(p, w)$ is the solution of the utility maximization problem based on a locally non-satiated and continuous preference $\succsim$. If there is no solution to this problem, then the preference was not representable in the first place. We will assume that the solution exists.

Definition: For each $\bar{u} \in \mathbb{R}$, define the upper contour set of $\bar{u}$ as $V_{\bar{u}} \equiv \{z \in \mathbb{R}^m_+ \mid u(z) \geq \bar{u}\}$, then $\forall p \in \mathbb{R}^m_{++}$, $e(p, \bar{u}) = \min\limits_{z \in V_{\bar{u}}} pz'$ so

$$V_{\bar{u}} \subseteq \bigcap_{p \in \mathbb{R}^m_{++}} \{z \in \mathbb{R}^m_+ \mid pz^\top \geq e(p, \bar{u})\}$$

Proof: Recovering the Utility Function from the Expenditure Function

Given how we can recover the indirect utility function u from the expenditure function e, and using the fact that e and the indirect utility function v are inverse functions of each other (i.e. $w = e(p; u)$) can be inverted to obtain $u(p, w) = e^{-1}(p, w) = v(p, w)$ and vice versa), show that you can obtain e and u from v.

By Roy's identity and duality, we know that

$$\frac{\partial v(p, w)/\partial p_k}{\partial v(p, w)/\partial w} = x_k(p, w) = h_k(p, v(p, w)) = \frac{\partial e(p, v)}{\partial p_k}$$

Let $u(x(p,\bar{w})) = \bar{u}$ and $\bar{w} = e(p,\bar{u})$. At $(\bar{w},\bar{u})$, we know that the Hessian matrix of the expenditure function with respect to prices is equal to the Slutsky's substitution matrix and it is negative semi-definite. This means that we have a system of m^2 partial differential equation's to solve. If we integrate with respect to prices, we get the first order conditions between the expenditure function and the indirect utility function.

Given the condition that $\bar{w} = e(p,\bar{u})$, we can thus locally trace out the expenditure function at each $(w,u) \in \mathbb{R}_+ \times \mathbb{R}$ by observing how the expenditure moves along with u and thus recover the expenditure function.

After recovering the expenditure function, we can attempt to recover the utility function by realizing that since $v(p,w)$ and $e(p,\bar{u})$ are both optimized functions, we have

$$\forall e_1, e_2 \in \mathbb{R}_{++},\, e_1 \geq e_2 \iff u(x(p,e_1)) \geq u(x(p,e_2))$$

(if not, the consumer choice will violate C-WARP). We can thus construct a "preferred" set for each utility level $\bar{u}$ and prices p as:

$$X(\bar{u}) \equiv \{x \in \mathbb{R}^m_+ \mid u(x) \geq \bar{u}\} \underbrace{=}_{\text{From the } \iff \text{ above}} \{z \in \mathbb{R}^m_+ \mid pz^\top \geq e(p,\bar{u})\}$$

Moreover, since this set containment relationship must work for all prices, we can rewrite it as:

$$X(\bar{u}) = \bigcap_{p \in \mathbb{R}^m_+} \{z \in \mathbb{R}^m_{++} \mid pz \geq e(p,\bar{u})\}$$

Notice that, assuming continuous, locally non-satiated, and monotonic preferences. Define $\{0\} = \{x \in \mathbb{R}^m_+ \mid u(x) = 0\}$, then $\forall \bar{u} \in \mathbb{R}_{++}$, $X(\bar{u})$ is

- Convex: $\forall \alpha \in (0,1), x, y \in X(\bar{u})$,

$$\alpha px \geq \alpha e(p,\bar{u}),\, (1-\alpha)py \geq (1-\alpha)e(p,\bar{u}) \iff p[\alpha x + (1-\alpha)y] \geq e(p,\bar{u})$$

- Closed: $\forall (x_i)_{i=1}^\infty \in X(\bar{u})$ such that $x_i \to x$,

$$px_i \geq e(p,\bar{u}) \;\Rightarrow\; \lim_{i\to\infty} px_i \geq \lim_{i\to\infty} e(p,\bar{u}) \equiv e(p,\bar{u} \;\Rightarrow\; px \geq e(p\bar{u}) \;\Rightarrow\; x \in X(\bar{u})$$

- Bounded below: $\forall \bar{u} \in \mathbb{R}_{++}$, $\exists \underline{x} \in \mathbb{R}^m_+ \setminus X(\bar{u})$ such that $\forall x \in X(\bar{u}), x >> \underline{x}$.

  Proof: Suppose not, then $\forall \underline{x} \in \mathbb{R}^m_{++} \setminus X(\bar{u})$, $\forall x \in X(\bar{u})$, $\underline{x} \succsim x$, so $p\underline{x} \geq px$, meaning $\underline{x} \in X(\bar{u})$, which is a contradiction to the assumption that $\underline{x} \in \mathbb{R}^m_+ \setminus X(\bar{u})$.

This means that we can further define the "indifference" set for each $\bar{u} \in \mathbb{R}_+$ as:

$$\mathbb{X}(\bar{u}) \equiv \{x \in \mathbb{R}^m_+ \mid u(x) = \bar{u}\}$$

Notice that since the "preferred" sets $X(\bar{u})$ are convex, closed, and bounded below for all $\bar{u} \in \mathbb{R}_+$, it means that $\mathbb{X}(\bar{u})$ is convex, compact[a], and most importantly, creates a partition of $\mathbb{R}^m_+$ based on different $\bar{u} \in \mathbb{R}_+$.

Hence $\mathbb{X}(u)$ describes a surjective and one-to-many mapping (a.k.a. a correspondence) from $\mathbb{R}_+$ to $\mathbb{R}^m_+$. The "inverse" of this correspondence is then the utility function of the consumer recovered from the expenditure function.

□

[a]If this is not intuitive to you, think about this like slicing a ham. The set $X(\tilde{u})$ is the ham, and it is convex, compact, and bounded below. Thus, each slice of ham $\mathbb{X}(\tilde{u})$ is also convex and compact, and the now disjoint slices together form partitions that make up the ham.

8.3 Welfare Evaluation of Price Changes

When there the prices of goods change, either via a stochastic shock or policy changes, we, as economists, should ask 2 questions: Are the consumers better or worse off under the new prices? How much better or worse off are they compared to the old prices?

Recall that weel-behaved consumers consume $x(p, w)$ such that $px^\top(p, w) \leq w$. Suppose now that there is a price change from p to p' and the consumer consumes at a new bundle $x(p', w)$ such that $p'x(p', w) \leq w$ and $x(p', w) \neq x(p, w)$.

By C-WARP, this means either $p'x(p, w) > w$ or $px(p', w) > w$. Luckily, we have the great tool - *indirect utility function* $v : \mathbb{R}^m_{++} \times \mathbb{R}_+ \to \mathbb{R}$ that is derived from the preference $\succsim$ to use for comparison. Since v is the inverse of e, this means we can evaluate changes welfare using the changes in expenditure.

Specifically, for any price changes from p to p', we can take an arbitrary price vector $\bar{p}$ and compare $e(\bar{p}, v(p, w))$ and $e(\bar{p}, v(p', w))$.

Definition: The **Equivalent Variation** is how much money a social planner can charge a consumer in order for the consumer to stay at the old prices p.

$$EV(p, p', w) \equiv e(p, v(p', w)) - e(p, v(p, w)) = e(p, v(p', w)) - w$$

Definition: The **Compensating Variation** is how much money a social planner needs to compensate the consumer in order to move to new prices p'.

$$CV(p, p', w) \equiv e(p', v(p', w)) - e(p', v(p, w)) = w - e(p', v(p, w))$$

Equivalently, we can write

$$\begin{aligned} e(p, v(p, w + EV)) &= e(p, v(p', w)) \\ e(p', v(p', w - CV)) &= e(p', v(p, w)) \end{aligned}$$

8.4 Aggregate Demand

So far, what we have studied is simply the consumption behaviors of individual households. But that does not necessarily mean that we can scale up our findings in aggregate. As a thought experiment, think about a Giffen good for a given household. Would it even make sense that it is also a Giffen good for the entire world? This absurd-ish example shows us that our findings are not necessarily scale-able in aggregate, and we have to make some pretty strong assumptions in order to try.

Let $\mathcal{I}$ be the set of $|\mathcal{I}| = N \in \mathbb{N}$ agents. Take agents $i, h \in \mathcal{I}$, the wealth of agent i to be $w^i \in \mathbb{R}_+$, and the set of goods $M = \mathbb{R}^m_+$. Let $w = (w^1, \ldots, w^n) \in \mathbb{R}^n_+$ denote the vector of wealth for all agents. Let $W^N = \sum_{i=1}^{n} w^i$ denote the aggregate wealth of all agents in N. Let the demand correspondence of agent i be $x^i : \mathbb{R}^m_{++} \times \mathbb{R}_+ \rightrightarrows \mathbb{R}^m_+$ and the aggregate demand correspondence be $x : \mathbb{R}^m_{++} \times \mathbb{R}_+ \rightrightarrows \mathbb{R}^m_+$ such that $x(p, w) = \sum_{i=1}^{n} x^i(p, w^i)$.

Empirically, the aggregate demand/consumption correspondence x depends only on the aggregate wealth w^N and not the wealth vector $w \in \mathbb{R}^n_+$. Thus, for us to do analysis on aggregate demand, we would need to assume that $\forall w, \tilde{w} \in \mathbb{R}^n_+ \; s.t. \; w^n = \tilde{w}^N$, $x(p, w) = x(p, \tilde{w})$. This assumption, among other things, is equivalent to saying $\forall i, j \in \mathcal{I}$

$$\frac{\partial x^i_k(p, w^i)}{\partial w} = \frac{\partial x^j_k(p, w^j)}{\partial w}$$

meaning that the income effect on a good k should be identical between all agents. Clearly, this is highly improbable (think about our very simplistic example with the Giffen goods).

If we try to relax this assumption by just a little bit, then we end up with assuming that all agents/households have homothetic preferences. That is, however, still a very strong assumption as it in turn assumes somewhat homogeneous preferences.

If we try to relax this just a little bit more, it means that all agents at least have quasi-linear preferences in money. This is a little bit more believable, but it is still relatively easy to come up with counter-examples. For the remainder of this chapter, let's assume this less absurd assumption and see how far we can go.

Theorem 8.2: SMD (Sonneschein (1973), Mantel (1974), and Debreu (1974))

Let ω denote a matrix of endowment of goods in M such that $\omega = (\omega^1, \omega^2, \dots, \omega^n) \in \mathbb{R}^{m\times n}_{++}$ such that $p \cdot x^i(p, \omega^i) \leq p \cdot \omega^i$ and $\omega^N = \sum\limits_{i\in\mathcal{I}} \omega^i$

Suppose that $\succsim^i$ is non-decreasing.

Let $f : \mathbb{R}^m_{++} \times \mathbb{R}^{m\times n}_{++} \to \mathbb{R}^m_+$ be a continuous function that is homogeneous of degree 0 in (p, w) such that $p \cdot f(p, \omega) = p \cdot \omega^N$.

Then, $\forall \varepsilon \in \mathbb{R}_{++}$, $\exists n \in \mathbb{N}$ consumers with preference $\succsim^i$ that are continuous, non-decreasing, and strictly convex and $\exists \omega \in \mathbb{R}^{m\times n}_{++}$ such that $x : \mathbb{R}^m_{++} \times \mathbb{R}^{m\times n}_{++}$ satisfies $\forall p \in \mathbb{R}^m_{++}$, $x(p, \omega) = f(p, \omega)$ such that $\forall k \in M$, $p_k > \varepsilon$.

Note: Readers should be aware that the aggregate demand function need not satisfy *C-WARP*, even if all consumers in $\mathcal{I}$ do. This means that the aggregate choices may not be rationalizable.

Proposition (MWG 4.B.1): The necessary and sufficient condition for $x(p, w)$ depend only on w^N is[32]

$$\exists a^i(p) \in \mathbb{R}, b(p) \in \mathbb{R}_{++} \; s.t. \; v^i(p, w^i) = a^i(p) + b(p)w^i$$

[32] Recall that indirect utility functions are unique up to affine transformations.

Now, suppose that the aggregate demand correspondence satisfies continuity, local non-satiation, and strict convexity (so that x is a function) so that $x^i(p, w^i)$ and $x(p, w)$ are continuous functions that are homogeneous of degree 0 in $(p, w^i), (p, w)$.

Another approach to aggregate demand analyses is to assume that the society, in aggregate, behaves like a single agent/household, then the aggregate choices would be rationalizable.

Definition: We say that **positive representative agent** exists for a given aggregate demand $x : \mathbb{R}^m_{++} \times \mathbb{R}_+ \to \mathbb{R}^m_+$ if $\exists \succsim$ on $\mathbb{R}^m_+$ that rationalizes $x(p, w)$.

Definition: A **Bergson-Samuelson social welfare function** $W : \mathbb{R}^n \to \mathbb{R}$ is a function that maps the utility of all n consumers to an aggregate utility.

Definition: The **Social Planner's Problem** is an aggregate utility maximization problem described by

$$\max W(v^1(p, w^1), \ldots, v^n(p, w^n)) \ s.t. \ \sum_{i=1}^{n} w^i = \bar{w}^N, \ (w^1, \ldots, w^n) \in [0, \bar{w}^N]^n$$

The social indirect utility function is $v(p, w) = W(v^1(p, w^1(p, \bar{w}^N)), \ldots, v^n(p, w^n(p, \bar{w}^N)))$.

Remark: Readers should note that, by definition, any solution to the social planner's problem is Pareto optimal.

Proposition: The social indirect utility function $v : \mathbb{R}^m_{++} \times \mathbb{R}_+ \to \mathbb{R}$ is an indirect utility function of a positive representative agent for the aggregate demand $x : \mathbb{R}^m_{++} \times \mathbb{R}_+ \to \mathbb{R}^m_+$ given by $x(p, \bar{w}^N) = \sum_{i=1}^{n} x^i(p, w^i(p, \bar{w}^N))$.

Did you know that about 95% of the people in New York City use an iPhone?

They are big Apple people.

9 Producer Theory

Now that we have learned some methods of modeling consumer behavior, we should try to see if we can extend it to model producer behavior. First, we will study producers in competitive markets. Let $\mathcal{J}$ denote the set of producers such that $\|\mathcal{J}\| = J \in \mathbb{N}$. Let $j \in \mathcal{J}$ denote a producer. Define $y^j \in \mathbb{R}^m$ to be the vector of production for producer j. The production set is $Y^j \subseteq \mathbb{R}^m \;\Rightarrow\; y^j \in Y^j \subseteq \mathbb{R}^m$.

Definition: The **production/transformation function**[33]**.** of producer j is $F^j : \mathbb{R}^m \to \mathbb{R}$ that is strictly increasing.

Definition: The **production/transformation frontier** of j is $\{\tilde{y} \in Y^j \mid F^j(\tilde{y}) = 0\}$.

Definition: The **production possibility set** is $Y^j = \{\tilde{y} \in \mathbb{R}^m \mid F^j(\tilde{y}) \leq 0\}$. Define the production possibility set in the market as $\underbrace{Y \equiv \prod_{j \in \mathcal{J}} Y^j}_{\text{Cartesian Product}}$, the transformation vector $F \equiv (F^1, \dots F^j)$, and the production vector $y \equiv (y^1, \dots, y^J)$

Definition: Assume that f^j is differentiable and $F(\tilde{y}) = 0$, we define the **Marginal Rate of Transformation** as:

$$MRT_{k,l}(\tilde{y}) \equiv \frac{\partial F(\tilde{y})/\partial \tilde{y}_k}{\partial F(\tilde{y})/\partial \tilde{y}_l}$$

Properties that Y^j can have: (*We will always assume the first four properties*)

(i) **No Free Production**: $Y^j \cap \mathbb{R}^m_+ = \{0\}$

(ii) **Free Disposal**: $\hat{y} \in Y^j \wedge \tilde{y} \leq \hat{y} \;\Rightarrow\; \tilde{y} \in Y^j$

(iii) **Irreversibilty**: $\tilde{y} \in Y^j \backslash \{0\} \;\Rightarrow\; -\tilde{y} \notin Y^j$

(iv) Y^j is **closed**.

——— The next 2 are not always assumed ———

(v) Constant Returns to Scale: Production frontier is linear within each orthant[34]

[33] I like to think about F here not as our standard production function, rather, the output should be thought of as waste. Essentially, if $F(y) = 0$, there is no "waste" (but not necessarily most efficient, because that would depend on prices of inputs and outputs)

[34] A multi-dimensional "quadrant".

(a) **Decreasing/Non-Increasing Returns to Scale**: Y^j is convex. Note that this means $\forall \tilde{y} \in Y^j$, $\lambda \in [0,1]$, we have $\lambda\tilde{y} \in Y^j$.

(b) **Increasing/Non-decreasing Returns to Scale**: $\forall \tilde{y} \in Y^j$, $\lambda \in [1,\infty)$, we have $\lambda\tilde{y} \in Y^j$.

(vi) **Recession Cone Property**: $\tilde{y} \in Y^j \wedge \tilde{y}_k > 0 \;\Rightarrow\; \exists\bar{\lambda} \in [1,\infty)\, s.t. \forall\lambda \in (\bar{\lambda},\infty)$, $\lambda\tilde{y} \notin Y^j$

Our typical producer problem is a profit maximization problem naively described as: For some $p \in \mathbb{R}^m_{++}$

$$\max_{\hat{y}\in\mathbb{R}^m} p\cdot\hat{y} \;\; s.t. \;\; F(\hat{y}) \leq 0$$

Just like the consumer problem, we can try to solve this with

$$\mathcal{L}(\hat{y},\lambda) = p\cdot\hat{y} - \lambda(F(\hat{y}))$$

giving us the first order conditions that a solution must satisfy:

$$\begin{aligned} \forall k \in M,\; p_k &= \lambda\frac{\partial F^j(\hat{y})}{\partial \hat{y}^j_k}\Big|_{\hat{y}=y^*} \\ F(y*) &= 0 \end{aligned}$$

Recall (Sundaram 7.16, 8.13) that if the <u>domain/constraint set is convex</u> and the <u>objective function is concave</u>, then the Kuhn-Tucker first order conditions are necessary and sufficient for y^* to be a maximizer.

Alternatively, if the <u>domain is convex</u> and the objective function is <u>*quasi-concave*</u>, AND $\nabla f \neq 0$, then the Kuhn-Tucker conditions are necessary and sufficient.

Definition: The producer's **Supply Correspondence** is $y^j : \mathbb{R}_{++} \rightrightarrows Y^j$ such that

$$y^j(p) \equiv \operatorname*{argmax}_{\hat{y}\in Y^j} p\cdot\hat{y}$$

Definition: The producer's profit function is $\Pi^j : \mathbb{R}_{++} \to \mathbb{R}$ such that

$$\Pi^j(p) = \max_{\hat{y}\in Y^j} p\cdot\hat{y}$$

Definition: Similar to the consumer problem, at the solution y^*, we have $\forall k, l \in M$

$$MRT_{k,l} = \frac{p_k}{p_l}$$

Proposition: Let Y^j satisfy assumptions 1-4 and the recession cone property, then the producer's problem has a solution y^j that satisfies:

(i) y^j is homogeneous of degree 0 in prices p

(ii) $\forall p \in \mathbb{R}^m_{++}$, if Y^j is (strictly) convex, then y^j is a (function) convex-valued correspondence.

(iii) $\forall p \in \mathbb{R}^m_{++}$, if y^j is differentiable at p, then $Dy^j(p) = D^2\Pi^j(p)$ and $Dy^j(p)$ is symmetric and positive semi-definite[35].

Definition: Law of Supply says that $\forall p, p' \in \mathbb{R}^m_{++}$, $\hat{y} \in y^j(p)$, $\hat{y'} \in y^j(\hat{p})$ we have

$$(p - p')(\hat{y} - \hat{y'}) \geq 0$$

so firms always respond positively to prices. Note that this is exact opposite of the relationship as in Hicksian demand.

Proposition (MWG 5.C.1): Let $j \in \mathcal{J}$ and Y^j satisfy assumptions 1-4, convexity, and the recession cone property, then the profit function $\Pi^j(p)$ is

(i) Homogeneous of degree 1 in prices and convex

(ii) (**Hotelling's Lemma**) $\forall p \in \mathbb{R}^m_{++}$ such that $y^j(p)$ is a singleton, $\Pi^j(p)$ is differentiable at p and $\frac{\partial}{\partial p_k}\Pi^j(p) = y^j_k(p)$

Proposition (Kreps 9.9): Let $y^j(p)$ be the solution to the profit maximization problem. If y^j is locally bounded and upper-hemi-continuous, then $\Pi^j(p)$ is continuous in p.

9.1 Simple Producers with Single Output

Now we will consider the case where we have multiple inputs and 1 output ($m \in M$). Take output $q^j \in \mathbb{R}_+$, inputs $d^j \in \mathbb{R}^{m-1}_+$ $s.t.$ $q^j = f^j(d^j)$. The new production possibility set is $Y^j \equiv \{(-d, \tilde{q}) \in \mathbb{R}^m \mid \tilde{q} \leq f^j(d) = q\}$.

[35] Comes from the law of supply

Definition: The ***q*-level isoquant** is defined as $iso(q) \equiv \{d^j \in \mathbb{R}_+^{m-1} \mid f^j(d) = q\}$

Definition: The marginal rate of substitution of inputs d^j on the $q-$level isoquant is called the **Marginal Rate of Technical Substitution**, which is defined as:

$$MRTS_{k,l}^j(d) = \frac{\partial f / \partial d_k^j}{\partial f / \partial d_l^j}(d)$$

Let p_m be the price of output and p_{-m} be the prices of inputs, we can rewrite the profit maximization problem as:

$$\max_{d \in \mathbb{R}_+^{m-1}} p_m f^j(d) - p_{-m} d$$

At the solution d^*, we have that $MRTS_{k,l}(d^*) = \frac{p_k}{p_l}$. From this problem, we get $d^j(p)$ as the **input demand function** and $q^j(p) = f^j(d^j(p))$ as the **supply function**. This allows us to consider a cost minimization version of the producer's problem given a output level $\bar{q}$.

Formally, the problem is equivalent to:

$$\min_{d \in \mathbb{R}_+^{m-1}} p_{-m} d \ s.t. \ f^j(d) \geq \bar{q}$$

which can be solved with the Lagrangian:

$$\mathcal{L}(d, \lambda) = \sum_{k=1}^{m-1} p_k d_k - \lambda(f^j(d) - \bar{q})$$

with the necessary first order conditions for the solution d^*:

$$\forall k \in \mathbb{R}^{m-1}, \ \frac{\partial \mathcal{L}}{\partial p_k}(d^*) = 0$$
$$MRTS_{kl}^j(d) \equiv \frac{\partial f^j / \partial d_k}{\partial f / \partial d_l}(d^*) = \frac{p_k}{p_l}$$

Definition: The **Cost Function** of producer j is $c^j(p_{-m}, \bar{q}) : \mathbb{R}_{++}^{m-1} \times \mathbb{R}_+ \to \mathbb{R}_+$.

Definition: The **Profit Function** of producer j is $\Pi^j : \mathbb{R}_{++}^m \to \mathbb{R}$ is defined as:

$$\Pi^j(\bar{q}, p) = \max_{\bar{q} \in \mathbb{R}_+} p_m \bar{q} - c^j(p_{-m}, \bar{q})$$

Proposition: The input demand correspondence $d^j : \mathbb{R}^{m-1}_{++} \times \mathbb{R}_+ \rightrightarrows \mathbb{R}^{m-1}_+$ satisfies:

(i) Homogeneous of degree 0 in prices p

(ii) If $f^j(d)$ is quasi-concave, then $d^j(p_{-m}, \bar{q})$ is convex(-valued).

(iii) If $f^j(d)$ is strictly quasi-concave, then $d^j(p_{-m}, \bar{q})$ is a singleton/function of (p_{-m}, q).

Proposition: The cost function $c^j : \mathbb{R}^{m-1}_{++} \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies

(i) Homogeneous of degree 1 in prices p

(ii) c^j is concave in prices p

(iii) c^j is non-decreasing in q

(iv) If Y^j is convex (meaning non-increasing returns to scale), then c^j is convex in $\bar{q}$

(v) (**Shepard's Lemma**) $\forall (\hat{p_{-m}}, \hat{q}) \in \mathbb{R}^{m-1}_{++} \times \mathbb{R}_+$, if $d^j(\hat{p_{-m}}, \hat{q})$ is a singleton, then c^j is differentiable with respect to p_{-m} at $(\hat{p_{-m}}, \hat{q})$ and

$$d^j_k(\hat{p_{-m}}, \hat{q}) = \frac{\partial c^j(\hat{p_{-m}}, \hat{q})}{\partial p_k}$$

If $d^j_k(\hat{p_{-m}}, \hat{q})$ is further differentiable in p_{-m}, then $D_{p_{-m}} d^j(\hat{p_{-m}}, \hat{q})$ is symmetric and negative-semi-definite.

Readers are encouraged to consult MWG 5.D for more graphical intuitions.

9.2 Aggregate Supply

Let Y^j denote the production possibility set of producer j. The production possibility set of the economy is defined as $Y = \prod\limits_{j \in \mathcal{J}} Y^j$. Irreversibility still applies in aggregate; formally, $\forall \hat{y} \in \mathbb{R}^m \setminus \{0\}$, if $\exists h \in \mathcal{J}$ such that $\hat{y} \in Y^h$, then $\{(\hat{y}^1, \hat{y}^2, \ldots, \hat{y}^J) \in \prod\limits_{j \in \mathcal{J}} \mid \sum\limits_{j \in \mathcal{J}} \hat{y}^j = -\hat{y}\} = \varnothing$. i.e., all vectors $\hat{y}^j$ and $\hat{y}$ are linearly independent.

Definition: The **Aggregate Supply Correspondence** $y : \mathbb{R}^m_{++} \rightrightarrows \sum\limits_{j \in \mathcal{J}} Y^j$ is $y(p) = \sum\limits_{j \in \mathcal{J}} y^j(p)$. Additionally, we know that

$$\nabla_p y^j(p) = \begin{pmatrix} \nabla_p y^j_1(p) \\ \vdots \\ \nabla_p y^j_m(p) \end{pmatrix} = \begin{pmatrix} \frac{\partial y^j_1(p)}{\partial p_1} & \cdots & \frac{\partial y^j_1(p)}{\partial p_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial y^j_m(p)}{\partial p_1} & \cdots & \frac{\partial y^j_m(p)}{\partial p_m} \end{pmatrix}$$

is *symmetric* and *positive semi-definite* and $y(p)$ is *homogeneous of degree 0* in p. This makes sense because the inputs are *produced* in negative values, so the *diagonal* entries in the matrix are not the own-price effect of inputs and outputs, but rather, the own-price effects of outputs and *negative inputs*.

Definition: $y^* : \mathbb{R}^m_{++} \rightrightarrows \sum\limits_{j \in \mathcal{J}} Y^j$ is the solution to the profit maximization problem of the **Aggregate producer**. $\Pi^* : \mathbb{R}^m_{++} \to \mathbb{R}$ is the profit function of the aggregate producer.

Definition: $\forall \hat{y} \in \sum\limits_{j \in \mathcal{J}} Y^j$, if $\hat{y}$ is profit-maximizing over $\sum\limits_{j \in \mathcal{J}} Y^j$ for the price vector $p \in \mathbb{R}^m_{++}$, then $\hat{y}$ is **Efficient**.

Definition: If $\sum\limits_{j \in \mathcal{J}} Y^j$ is convex, then any efficient $\hat{y} \in \sum\limits_{j \in \mathcal{J}} Y^j$ is profit-maximizing for some $p \in \mathbb{R}^m_+$

Proposition: $\forall p \in \mathbb{R}^m_{++}$, $\Pi^*(p) = \sum\limits_{j \in \mathcal{J}} \Pi^j(p)$, $y^*(p) = \sum\limits_{j \in \mathcal{J}} y^j(p)$. So the aggregate optimal production is the sum of individual firms' aggregate production.

Since producers' do not face an (long-run) budget constraint, they face exactly the Hicksian problem. As such, there is no "income-effect for producers" and we need not worry about how the aggregate producer behavior may be inconsistent with what we know about individual producers. This is why law of supply simply always holds like how law of demand always holds for Hicksian demand correspondences.

How do young bivalves settle their disputes?

They go to the small claims court.

10 General Equilibrium

Now that we have studied consumer and producer behaviors separately, it is time to see what happens when they interact in an economy. The idea of an equilibrium seems straightforward, but as we will come to find out, it is quite delicate. In essence, we want to describe not the process of trade, but the end result of trade. Later on, we will briefly discuss the stability of equilibria as well as their uniqueness, or the lack thereof. We shall start with a simple thought experiment, without restrictions of the market environment. As a student of economics, I believe nothing describes the nuance quite as well as the following excerpt from p.23 of *General Competitive Analysis* (Arrow and Hahn 1980):

> The decision to supply a good in a perfectly competitive economy is not a decision to supply [x amount of] goods to [n] agents, but simply to exchange [x amount] of the good for other goods. If the [price] for a good is zero and agents plan to supply some of it, then by the assumption of *free disposal*, we may simply say that agents decide to dispose of a certain amount of the good. Clearly this decision can be carried out by our assumption, whatever the demand of other agents may be. If this demand was greater than the amount offered, then the decisions of the demanding agents cannot be carried to fruition. From this, we conclude that while we would never be willing to regard a situation with positive excess demand in some market as an equilibrium, <u>an excess supply in a market where the price is zero is quite consistent with our notion of an equilibrium.</u> All this seems agreeable to common sense, and it remains to put it more formally.

For now, we will assume:

(1) Competitive market (all consumers and producers are price-takers).

(2) All agents have perfect information.

(3) All agents have perfect property rights.

(4) The economy has no frictions.

Formally, we will define an economy E that contains:

- N: The set of $n \in \mathbb{N}$ consumers.
- $\mathcal{J}$: The set of $J \in \mathbb{N}$ producers.
- M: The set of $m \in \mathbb{N}$ goods.

- $X^i \subseteq \mathbb{R}_+^m$: The consumption set of consumer i.
- $\omega_k^i \in \mathbb{R}_+$: Consumer i's endowment of good k.
- $\omega^i \in \mathbb{R}_+^m$: Consumer i's vector of endowments.
- $\omega \in \mathbb{R}_+^{m\times n}$: The endowment matrix of m goods of all n consumers.
- $\omega_k^N \equiv \sum\limits_{i=1}^{n} \omega_k^i$, $w^N \equiv (\omega_1^N, \ldots, \omega_m^N)$
- $\theta^i \equiv (\theta_1^i, \ldots, \theta_J^i)$: The "investment portfolio" of consumer i on all J producers.
- $\sum\limits_{i\in\mathcal{I}} \theta_j^i = 1$: Producers are completely owned by consumers.
- In equilibrium, the market clears: $\sum\limits_{i\in\mathcal{I}} x_k^i = \sum\limits_{i\in\mathbb{N}} \omega_k^i + \sum\limits_{j\in\mathcal{J}} y_k^j$

We will further assume that consumer i has a continuous preference $\succsim^i$ over X^i and is representable by the utility function $u^i(x)$. $X = \prod\limits_{i\in\mathcal{I}} X^i \subseteq \mathbb{R}^{m\times n}$, $Y^j \subseteq \mathbb{R}^m$, $Y \equiv \prod\limits_{j\in\mathcal{J}} Y^j$.

Definition: An economy E is defined as a space $E \equiv \big((X^i, \succsim^i, \omega^i, \theta^i)_{i=1}^n, (Y^j)_{j=1}^J\big)$.

Definition: An **Allocation** (x, y) is a matrix $(x^1, \ldots, x^n; y^1, \ldots, y^J) = X \times Y \subseteq \mathbb{R}^{m\times(n+J)}$.

Definition: An allocation (x, y) is said to be **Feasible** if $\sum\limits_{i\in\mathcal{I}} x^i \leq \omega^N + \sum\limits_{j\in\mathcal{J}} y^j$.

Definition: We say that the allocation (x, y) **Pareto dominates** $(\hat{x}, \hat{y})$ if $\exists i \in \mathcal{I}\, s.t.\, x^i \succ^i \hat{x}^i$ and $\forall i \in \mathcal{I}$, $x^i \succsim^i \hat{x}^i$.

Definition: A feasible allocation $(x, y) \in X \times Y$ is **Pareto Optimal/Efficient** (PE) if it is NOT Pareto dominated by ANY other feasible allocation.

We want a simple economy such that, in equilibrium, consumers maximize their utility and producers maximize their profit. Since consumers have complete ownership of the producers, the profits of the producers become the expenditure of the consumers.

Definition (CE/WEA) : In the economy E, we say that $(x^*, y^*) \in X \times Y$ is a **Competitive Equilibrium/Walrasian Equilibrium Allocation** if

(i) Producers maximize their profit: $y^{*j} \in \underset{y^j\in Y^j}{\operatorname{argmax}}\, p^* y^j$

(ii) Consumers maximize their utility: $\forall i \in \mathcal{I}$, $x^{*i} \succsim^i \hat{x}^i$ where
$x^{*i}, \hat{x}^i \in B(p^*, \omega^i, \theta^i) \equiv \{x^i \in X^i \mid p^* x^{i\top} \leq p^* \omega^{i\top} + p^* y^* \theta^{i\top}\}$

(iii) Market Clears[36]: $\sum_{i \in \mathcal{I}} x_k^{*i} \leq \sum_{i \in \mathcal{I}} \omega_k^i + \sum_{j \in \mathcal{J}} y_k^{*j}$

Notice that market clearing condition requires that consumption is feasible, but the first two conditions require profit/expenditure/wealth add up. Also note that competitive equilibrium does NOT imply competitive market.

10.1 Partial Equilibrium (Marshall 1920)/Money in Utility

Recall our discussion in aggregate demand that we can try to study aggregate consumer behavior by assuming that the consumers' preferences are quasi-linear in money/numeraire.

The Marshallian world formalizes this idea and develops thoughts on partial equilibria. John Hicks showed that once we assume *constant marginal utility in money* in the Marshallian world, all consumer demand discussions are simplified to the Hicksian approach as income effect can be removed from the equation.

Let $u^i(x^i, m) = m^i + \phi^i(x^i)$ where m^i is the numeraire good (money). Assume that $\frac{\partial u}{\partial x^i} > 0$, $\frac{\partial^2 u}{\partial (x^i)^2}, \phi(0) = 0$.

Assume that producer j can produce any $q^j = \tilde{q}$ at cost $c^j(\tilde{q})$ where $c^j(0) \geq 0$ is strictly increasing and $\frac{\partial^2 c^j}{\partial x^2} \geq 0$.

Assume that aggregate demand $x : \mathbb{R}_{++} \to \mathbb{R}$ decreases in prices (law of demand holds in aggregate).

Assume that aggregate supply $q : \mathbb{R}_{++} \to \mathbb{R}$ increases in prices (law of supply holds in aggregate).

In equilibrium, we must have $x(p) = q(p) + \omega^N$.

Proposition: If (x^*, q^*) is a competitive equilibrium, supported by price p^*, in a Marshallian economy, then (x^*, q^*) is Pareto Efficient.

[36] Note that the inequality becomes equality if all consumer preferences are strongly monotonic or locally non-satiated.

Definition (Numeraire): The standard/base good for comparison.

> "We shall find it convenient, when dealing with multiple exchange, always to take some particular commodity as a standard of value (footnote: Numeraire, as Walras called it)." - J.R.Hicks, *Value and Capital*

Proposition: For any Pareto optimal utility levels, there exists a transfer of the numeraire good such that a competitive equilibrium of the Marshall economy after transfers attains the desired utility level.

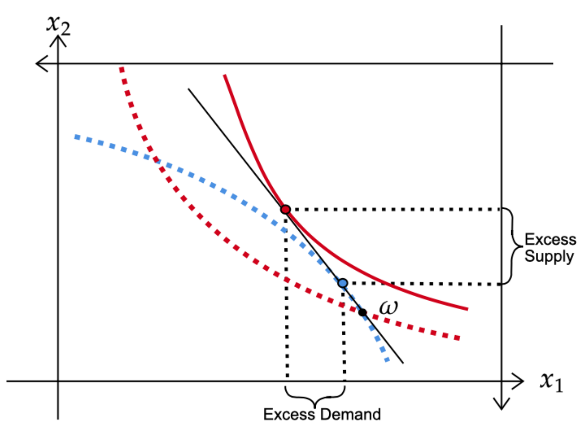
10.2 Exchange Economy (Edgeworth Box)

Consider the following Edgeworth box, where we have 2 well-behaved consumers with endowments in two goods and they trade in the economy.

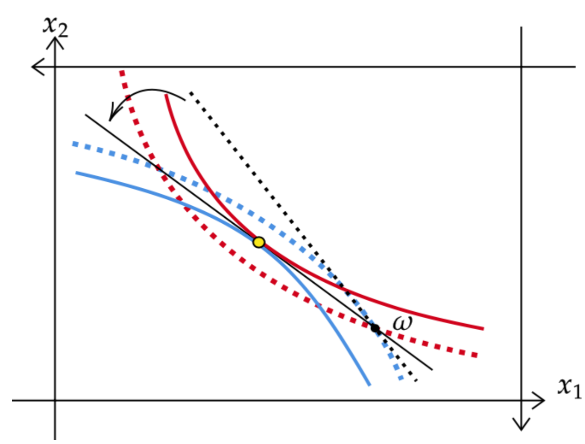
The consumption graph for consumer 2 is rotated so that the area between the two indifference curves are improvements that can be made to their consumption through trade.

Note that all allocation within the *Edgeworth box* are feasible as the box is drawn by total endowments. Also notice that any movement from the endowment point to inside the "leaf" is a Pareto improvement.

The consumers start out with the endowment ω and there is a price vector (black line) for trade but the market won't clear due to the excess demand of good 1.

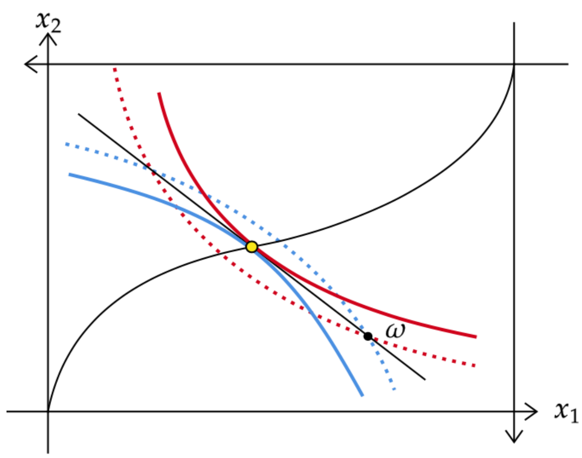


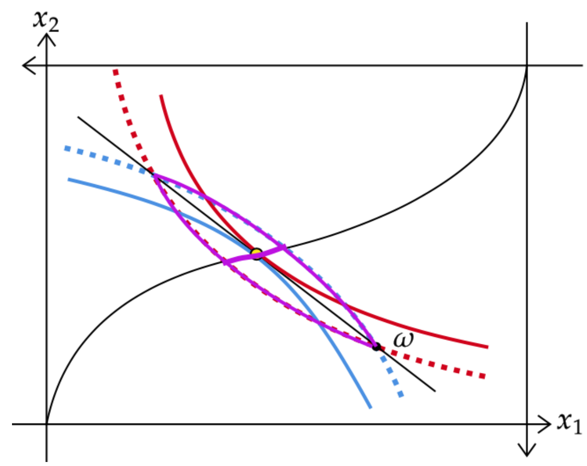
Through trade, the price vector will rotate from the dotted line to solid, and reach a new market clearing equilibrium with both consumers' utility level improved.



If we do this with some arbitrarily large amount of variaous endowments, we can trace a line through these equilibria, this line is called the **Pareto set**.

Any allocation not in the Pareto set but is in the purple leaf allows for both consumers to be better off by trading onto the Pareto set. The curve representing the intersection between the purple leaf and the Pareto set is called the **contract curve**.





10.3 Single Agent (Robinson-Crusoe) Economy

Consider a single-agent economy where the agent is endowed with 1 unit of time and 0 units of good c (i.e., $\omega = (1, 0)$). Suppose that the agent needs to split the time between leisure l and labor n so that $l + n = 1$.

Suppose that the agent can produce good c with a continuous, strictly increasing, strictly concave, and differentiable production function $c = q(n)$. Let the price of good c be 1 (good c is numeraire) and so the consumer gets "paid" in wages p_n for the labor they expend in producing good c.

We can thus formalize the agent's problem as:

- Profit: $q(n) - n \cdot p_n$
- Budget: $c = np_n + q(n) - np_n = q(n)$

- Lagrangian:

$$\begin{aligned}
\mathcal{L}(l,c,\lambda) &= u(l,c) - \lambda(c - q(1-l)) \\
\frac{\partial \mathcal{L}}{\partial l} &= u_l - \lambda(-q_n \cdot (-1)) \;\Rightarrow\; \lambda q_n = u_l \\
\frac{\partial \mathcal{L}}{\partial c} &= u_c - \lambda(1) \;\Rightarrow\; \lambda = u_c \\
\Rightarrow \;\frac{u_l}{u_c} &= q_n
\end{aligned}$$

From the maximization problem we get the supply of labor n^s as a function of u and q.
At optimal production, $q_n = p_n$, so the demand of labor n^d is a function of wage p_n.
At optimal consumption, $\frac{u_l}{u_c} = \frac{p_l}{p_c} = -p_n$.
In equilibrium, we have $n^s = n^d$, $c = q(n^s)$, and $p_l = q_n = -p_n$.

10.4 Welfare Theorems

Theorem 10.1: First Welfare Theorem

Let E be an economy where the consumers have *locally non-satiated preferences* and $(x^*, y^*) \in \mathbb{R}^m_+ \times \mathbb{R}^m_+$ is a *competitive equilibrium* supported by $p^* \in \mathbb{R}^m_{++}$, then (x^*, y^*) is *Pareto efficient*.

Proof 10.1: First Welfare Theorem

Let (x^*, y^*, p^*) be a competitive equilibrium. Suppose otherwise that (x^*, y^*, p^*) is not Pareto efficient, then there exists some other **feasible** allocation (x, y, p^*) that Pareto dominates (x^*, y^*, p^*).

Step 1: Since (x, y, p^*) Pareto dominates (x^*, y^*, p^*), we know that
$\exists i \in N$ such that $x^i \succ^i x^{*i}$
and
$\forall h \in N$, $x^h \succsim^h x^{*h}$.

Step 2: Since (x^*, y^*, p^*) is a competitive equilibrium, this must mean that
$p^* x^i > p^* \omega^i + \sum_{j \in \mathcal{J}} \theta^i p^* y^{*j}$

and
$p^*x^h \geq p^*\omega^h + \sum_{j\in\mathcal{J}} \theta^h p^* y^{*j}$.

Step 3: Since (x^*, y^*, p^*) is a competitive equilibrium, this must mean that the producer profit is maximized: $\forall j \in \mathcal{J}, y^j \in Y^j$, $p^*y^{*j} \geq p^*y^j$ and so

$$\sum_{i\in\mathcal{I}} p^*x^i > \sum_{i\in\mathcal{I}} p^*\omega^i + \sum_{j\in\mathcal{J}} p^*y^{*j} \geq \sum_{i\in\mathcal{I}} p^*\omega^i + \sum_{j\in\mathcal{J}} p^*y^j$$

But this means that (x, y, p^*) is not a feasible allocation. By contradiction, (x^*, y^*, p^*) must be the Pareto efficient allocation.

□

Definition: In an economy E, (x^*, y^*, p^*) is a **Competitive Equilibrium with Transfers** is there exists a vector $T \neq \vec{0}$ of *zero-sum wealth transfers* where $\sum_{i=1}^{n} T^i = 0$ such that

(i) Producers maximize their profit: $y^{*j} \in \operatorname*{argmax}_{y^j \in Y^j} p^*y^j$

(ii) Consumers maximize their utility: $\forall i \in \mathcal{I}$, $x^{*i} \succsim^i \hat{x}^i$ where

$$x^{*i}, \hat{x}^i \in B(p^*, \omega^i, \theta^i, T^i) \equiv \{x^i \in X^i \mid p^*x^{i'} \leq p^*\omega^{i'} + p^*y^*\theta^{i'} + T^i\}$$

(iii) Market Clears: $\sum_{i\in\mathcal{I}} x_k^{*i} \leq \sum_{i\in\mathbb{N}} \omega_k^i + \sum_{j\in\mathcal{J}} y_k^{*j}$

Theorem 10.2: Second Welfare Theorem

Assume that consumer preferences are continuous, convex, and strongly monotonic. Assume that Y^j satisfies assumptions 1-4, convexity, and the recession cone property. Assume that $\exists \tilde{y} \in Y$ such that $\omega + \tilde{y} \in \mathbb{R}^m_{++}$, then

$$(\hat{x}, \hat{y}) \text{ is Pareto efficient } \Rightarrow \exists p^* \in \mathbb{R}^m_{++}\, s.t.\, (\hat{x}, \hat{y}, p^*) \text{ is a CE with transfers } T \in \mathbb{R}^n$$

Proof 10.2: Second Welfare Theorem

Let $(\hat{x}, \hat{y})$ denote a Pareto optimal allocation in an economy $\hat{E} \equiv (\succsim^i, x^i, \omega^i, Y^j)$ where consumer preferences are continuous, convex, and strongly monotonic, $\forall j \in \mathcal{J}$, Y^j satisfies assumptions 1-4, convexity, and the recession cone property. Assume that $\exists \tilde{y} \in Y$ such that $\omega + \tilde{y} \in \mathbb{R}^m_{++}$. Consider a new production economy $E = (\succsim^i, x^i, \hat{x}^i, Y^j - \{\hat{y}^j\})$.

Since the consumers and producers are well-behaving, we know that a Walrasian Equilibrium (x^*, y^*, p^*) exists in E and we can calculate the p^* for the equilibrium.

Notice that since $(\hat{x}, \hat{y})$ is Pareto Optimal in $\hat{E}$, we know that

$$\sum_{i \in \mathcal{I}} \hat{x}^i = \sum_{j \in \mathcal{J}} \hat{y}^j + \sum_{i \in \mathcal{I}} \omega^i \;\Rightarrow\; \sum_{i \in \mathcal{I}} \hat{x}^i - \sum_{j \in \mathcal{J}} \hat{y}^j = \sum_{i \in \mathcal{I}} \omega^i$$

Since (x^*, y^*, p^*) is WEA in E, we must have that $x^{*i} \succsim^i \hat{x}^i$ and x^{*i} is feasible, so for some $\tilde{y}^j \in Y^j$, we have

$$\begin{aligned} \sum_{i \in \mathcal{I}} x^{*i} &\leq \sum_{j \in \mathcal{J}} y^{*j} + \sum_{i \in \mathcal{I}} \hat{x}^i = \sum_{j \in \mathcal{J}} (\tilde{y}^j - \hat{y}^j) + \sum_{i \in \mathcal{I}} \hat{x}^i \\ &= \sum_{j \in \mathcal{J}} \tilde{y}^j + \sum_{i \in \mathcal{I}} \hat{x}^i - \sum_{j \in \mathcal{J}} \hat{y}^j = \sum_{j \in \mathcal{J}} \tilde{y}^j + \sum_{i \in \mathcal{I}} \omega^i \end{aligned}$$

By our assumption, $(x^*, \tilde{y})$ is feasible in the original economy $\hat{E}$. Since $(\hat{x}, \hat{y})$ is Pareto optimal in $\hat{E}$, it must then mean that $\forall i \in \mathcal{I}$, $x^{*i} \sim^i \hat{x}^i$.

Since (x^*, y^*, p^*) is WEA and $0 \in Y^j - \{\hat{y}^j\}$, $\forall j \in \mathcal{J}$, $\Pi^j(y^j) = p^* y^{*j} \geq 0$, and so

$$\forall i \in \mathcal{I},\; \sum_{j \in \mathcal{J}} \theta^i_j \Pi^j(y^{*j}) + p^* \hat{x}^i \geq p^* \hat{x}^i$$

Since $\succsim^i$ are strongly monotonic and $x^{*i} \sim^i \hat{x}^i$, it must be that

$$\forall i \in \mathcal{I},\; p^* \hat{x}^i = p^* \hat{x}^i + \sum_{j \in J} \theta^i_j \Pi^j(y^{*j})$$

meaning that $\forall i \in \mathcal{I}$, $\sum_{j \in J} \theta_j^i \Pi^j(y^{*j}) = 0$, that is, all producers make 0 profit. So we have,

$$0 = p^* y^* = p^*(\tilde{y}^j - \hat{y}^j) \;\Rightarrow\; p^* \tilde{y}^j = p^* \hat{y}^j$$

Finally, define $T \in \mathbb{R}^n$ such that

$$\forall i \in \mathcal{I},\, p^* \hat{x}^i = p^* \hat{\omega}^i + \sum_{j \in \mathcal{J}} \left(\theta_j^i p^* \hat{y}^j\right) + T_i$$

Summing across $i \in \mathcal{I}$, we get

$$\sum_{i \in \mathcal{I}} p^* \hat{x}^i = \sum_{i \in \mathcal{I}} p^* \omega^i + \underbrace{\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (\theta_j^i p^* \hat{y}^j)}_{= \sum_{i \in \mathbb{N}} 0 = 0} + \sum_{i \in \mathcal{I}} T_i$$

Since $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (\theta_j^i p^* \hat{y}^j) = 0$, then $\sum_{i \in \mathcal{I}} T_i$ must also equal to 0, so $(\hat{x}, \hat{y}, p^*)$ is a WEA. □

Theorem 10.3: Second Welfare Theorem for an Exchange Economy

(Adapted from 2016 SS Part I Q2(a)) Let E be an exchange economy where all consumers $i \in \mathcal{I} : |\mathcal{I}| = N$ have preferences such that there exists a Walrasian Equilibrium with strictly positive prices for any endowment vectors. Let the consumer preferences be strongly monotonic and strictly convex and each consumer is endowed with $\omega_i \in \mathbb{R}_+^m$.

Let x^* be a Pareto Efficient Allocation in this economy, then there exists a lump-sum transfers $t = (t^1, t^2, \ldots, t^N) \in \mathbb{R}^n$ where

$$\sum_{i \in \mathcal{I}} t^i = 0$$

such that x^* is a Walrasian Equilibrium Allocation with transfers.

Proof 10.3: Second Welfare Theorem for an Exchange Economy

By assumption, there exists an allocation $(\hat{x}, \hat{p})$ such that $(\hat{x}, \hat{p}, \omega)$ is a Walrasian Equilibrium Allocation.

If $\hat{x} = x^*$, then our proof is done and $t = (0, \ldots, 0)$.

If $\hat{x} \neq x^*$, then since x^* is Pareto Optimal, it must be that $\exists\, i, j \in \mathcal{I}$ such that

$$\hat{x}^i \succ x^{*i} \;\wedge\; \hat{x}^j \prec x^{*j}$$

Since $(\hat{x}, \hat{p}, \hat{\omega})$ is a WEA, we know that

$$\hat{p}\hat{x}^i > \hat{p}x^{*i} \;\wedge\; \hat{p}\hat{x}^j < \hat{p}x^{*j}$$

and

$$\hat{p}\hat{x}^i = \hat{p}\omega^i \;\wedge\; \hat{p}\hat{x}^j = \omega^j$$

We can thus define a vector t of transfers such that

$$\hat{p}\hat{x}^i + t^i = \hat{p}x^{*i} \;\wedge\; \hat{p}\hat{x}^j + t^j = \hat{p}x^{*j}$$

Notice this implies that

$$\sum_{i \in \mathcal{I}} t^i = \sum_{i \in \mathcal{I}} \hat{p}(x^{*i} - \hat{x}^i) = 0$$

Since consumers have strictly convex preferences and x^* is a PEA, it must mean that x^* is the vector of unique maximizers such that

$$\forall i \in \mathcal{I},\, x^{*i} = \underset{x}{argmax}\ u^i(x)\ s.t.\ \hat{p}x^i = \hat{p}\omega^i + t^i$$

As such, x^* is a WEA with transfers $\square$

10.5 Social Welfare and General Equilibria

As economists, our job is to study how the society (market) moves to allocate resources. Through the first and second welfare theorems, we now have tools to move between Pareto optimal allocations and Walrasian/Competitive equilibria. To do this, we need to find a way to measure welfare, and find some kind of "mapping" between an equilibrium allocation and its social welfare.

Recall that a *Walrasian Equilibrium Allocation* (WEA) is an allocation that, **through prices**, allocates resources/endowments such that:

(i) Equalizes marginal rates of substitution across consumers (also called *allocative efficiency*).

(ii) Equalizes marginal rates of transformation across producers.

(iii) Equalizes marginal rates of substitution and marginal rates of transformation.

Definition: Let $W : \mathbb{R}^{m\times n} \to \mathbb{R}$ be the *linear* **Social Welfare Function** defined as

$$W(x) = \sum_{i\in\mathcal{I}} \lambda^i u^i(x^i)$$

where λ^i is the weight that consumer i's utility takes in the measuring of social welfare[37]. If we further assume that $u^i(x^i)$ is concave and that production functions are convex, then *every Pareto optimal allocation maximizes $W(x)$ for some weight λ_x*.

Now that we have a way to map allocations to welfare, how can we systematically determine the efficiency of any allocation? Recall our edgeworth box exercise. The reason we knew that endowment and the original price vector is inefficient (can be improved) is through observing that there are excess demand and supply on different goods in the market. To generalize this idea for m goods, we need to define **excess demand** in the market.

Definition: The **Excess Demand** correspondence of consumer i is a correspondence $z^i : \mathbb{R}^m_{++} \times \mathbb{R}^m_{+} \rightrightarrows \mathbb{R}^m$ defined as

$$z^i(p, \omega^i) = x^i(p, \omega^i) - \omega^i - \sum_{j\in\mathcal{J}} \theta^i y^i_j$$

[37] For example, a benign social planner who wants every agent's utility to matter the same will have $\lambda^i = \frac{1}{n}$

Definition: The **Aggregate excess demand** correspondence $z^N : \mathbb{R}^m_{++} \times \mathbb{R}^m_+ \rightrightarrows \mathbb{R}^m_+$ in the economy is defined as

$$z^N(p, \omega) \equiv \sum_{i \in \mathcal{I}} z^i(p, \omega^i)$$

Properties of the Aggregate Excess Demand Correspondence:
Assume that consumers in the economy have preferences $\succsim^i$ that are continuous, monotonic[38], and *strictly convex*. Then z^N is a *function* with the following properties[39]:

(i) Continuous

(ii) Homogeneous of degree 0 in prices

(iii) Walras's law is satisfied: $\forall p \in \mathbb{R}^m_{++}$, $p \cdot z^N(p, \omega) = 0$

(iv) Take a sequence of prices $(p^t)_{t=1}^{\infty} \in \mathbb{R}^2{}_{+}$, if $p^t \to p$ such that the price of good k, $p_k = 0$ and the price of the other goods, $p_{-k} > 0$, then $z^N_k(p^t, \omega) \to \infty$

Naturally, if we are interested in the efficiency of the economy (not allocative but overall), we want to look for allocations where the aggregate excess demand $= 0$. Notice that this is not equivalent to the individual excess demand are all 0 since that would mean that all agents are consuming their endowments and there is no production.

10.6 Existence of Walrasian Equilibria

10.6.1 Strictly Positive Price Walrasian Equilibria

In case you have not guessed it, showing the existence of a Walrasian equilibrium is not a hard task given the right assumptions.
Consider an economy E such that:

(i) $\forall i \in \mathcal{I}$, $X^i \subseteq \mathbb{R}^m_+$ convex, $\omega^i \in int(X^i)$, and consumer i's preference $\succsim^i$ is continuous, locally non-satiated, and strictly convex.

(ii) $\forall j \in \mathcal{J}$, Y^j satisfying assumptions 1-4, convexity, and the recession cone property.

Then since consumers and producers are assumed to be well-behaving, we can expect there to be at least one WEA. Formally, there are many ways to prove existence of WEA in different environments. We will go through a simple case of a trade economy with $m = 2$.

[38]Note that in lecture, this is the definition Eqguia gave. In MWG, this is suppsoed to be strongly monotonic
[39]MWG has an extra property here that the excess demand functions is bounded below

Proof: Existence of WEA for Trade Economy with $m = 2$

Consider a trade economy with 2 goods that has

(i) A set of consumers $\mathcal{I}$ with continuous and locally non-satiated[a] preferences $\succsim^i$, endowments ω^i.

(ii) Prices $p \in \mathbb{R}^2{}_{+}$

Let $x^i : \mathbb{R}^2_{++} \times \mathbb{R}^2_{+} \rightrightarrows \mathbb{R}^2_{+}$ denote the demand correspondence for agent i in this economy. Let $z^i : \mathbb{R}^2_{++} \times \mathbb{R}^2_{+} \rightrightarrows \mathbb{R}^2_{+}$ denote the individual excess demand and $z^N = \sum_{i \in \mathcal{I}} z^i(p, \omega^i)$.

For simplicity, we will normalize the price of good 2 to $p_2 = 1$ and we will denote the price of good 1 as p. If a WEA indeed exists, then $\exists p \in \mathbb{R}_+$ such that $z_1^N(p, \omega) = z_2^N(p, \omega) = 0$.

Notice that by Walras's law, we have

$$p \cdot z_1^N(p, \omega) + z_2^N(p, \omega) = 0$$

so if $\exists p \in \mathbb{R}_{++}\ s.t.\ z_1^N(p, \omega)$=0, then $z_2^N(p, \omega) = 0$ is automatic.
Take a sequence of price $(p_t)_{t=1}^{\infty} \in \mathbb{R}_{++}$ such that $p_t \to 0$, then by the property of aggregate excess demand, we know that $z_1^N(p_t, \omega) \to \infty$.

Next, observe that as $p \to \infty$, the aggregate excess demand $z_1^N(p, \omega)$ will behave as if p is fixed and $p_2 \to 0$, meaning that as $p \to \infty$, we have $z_2^N(p, \omega) \to \infty$. By Walras' law, this must mean that $z_1^N(p, \omega) \to -\infty$ as $p \to \infty$.

This means that $\exists p', p'' \in \mathbb{R}_{++}$ such that $z_1^N(p', \omega) > 0 > z_2^N(p'', \omega)$.

Excitingly, we also know that $z_1^N(p\omega)$ is continuous in (p, ω), so by the *Intermediate Value Theorem*, $\exists p^* \in (p', p'')$ such that $z_1^N(p^*, \omega) = 0$, and there indeed exists a WEA (x^*, p^*) in this economy. $\square$

[a] Readers should note that this proof in MWG (Prop. 17.B.2) actually required strongly monotonic and strictly convex preferences (needed for the singularity of demand correspondences) as well as strictly positive endowments.

10.6.2 Price 0 Walrasian Equilibria and Scarcity

At the beginning of this chapter, I hinted at the existence of equilibria where the price of certain goods are 0. This brief discussion is to provide some conditions and perhaps intuition for such equilibrium to exist.

Recall that in a WEA, consumers must maximize their utility given their budget. If the the price of good k is 0 in equilibrium, it must be that the marginal utility[40] of good k, at the equilibrium allocation, is less than or equal to 0. Moreover, since agents must be maximizing in the WEA, their equilibrium level utility must not be lower than their endowment level utility. i.e., The equilibrium allocation must not be Pareto Dominated by the original endowment.

Hopefully, this makes intuitive sense to you. If the price of good k is 0 in equilibrium, it must be that getting more of good k cannot benefit anyone, otherwise, the consumer is not maximizing. Under the assumption of *free disposal of the consumers* (not the same as free disposal of the producers), this means that goods that are not scarce, or goods that can cause satiation (e.g., a perfect compliment good that is just too abundant in the market) are often subject to price 0 equilibria. For the sake of discussion, let's call these goods **price-0 goods**.

Consequently, if an agent is endowed with only price-0 goods, their equilibrium utility must be locally satiated, but it need not be 0. Consider an exchange economy with 2 goods and 2 agents and both agents are endowed with 2 unit of each good. Suppose that their utility functions are:

$$u^1(x,y) = \begin{cases} x+y & , \ (x,y) \in [0,1]^2 \\ 0 & , \ \text{otherwise} \end{cases}$$

$$u^2(x,y) = \begin{cases} x+y & , \ (x,y) \in [0,3]^2 \\ 0 & , \ \text{otherwise} \end{cases}$$

Then we can obtain a price-0 equilibrium at:

$$(x^1, y^1, x^2, y^2, p_1, p_2) = (1,1,3,3,0,0)$$
$$u^1(x^1,y^1) = 2, \ u^2(x^2,y^2) = 6$$

[40]If the utility function is not differentiable (e.g., perfect compliments), the condition is that $p_k = 0$ in equilibrium if and only if an increase in k, in <u>any arbitrary amount</u>, does not increase utility.

This may seem counter-intuitive since agent 1 is simply giving away their endowment. However, notice that this allocation does not violate the definition of an WEA in anyway, since both agents are maximizing and the allocation is certainly feasible given the endowment. If you are having a hard time imagining free disposal here, think of it as agent 1 throwing half of their endowments in the dumpster and agent 2 simply goes and pick them up.

10.7 Uniqueness of A Walrasian Equilibrium

Surprisingly, or perhaps quite the opposite, the uniqueness of WEA is difficult to obtain without highly restrictive assumptions. Since we characterize WEA with aggregate excess demand, that difficulty of aggregate demand naturally carried over. Frustratingly, because of such restrictions, not only is uniqueness difficult to find in the general case, we also don't have general characterizations of an unique WEA.

In this section, we will discuss what we can but most of the time we can only find local unique equilibria and the global one is almost impossible to get.

Definition: We say the the demand correspondence[41] $x^i(p, \omega)$ satisfies the **Gross Substitute Property** if $\forall k \in M, \forall p, p' \in \mathbb{R}^m_+$,

$$p'_k > p_k \wedge p'_{-k} = p_{-k} \;\Rightarrow\; x^i_{-k}(p', \omega^i) > x^i_{-k}(p, \omega^i)$$

Proposition: An exchange economy where all consumer demand functions satisfy the gross substitute property has an unique WEA.

Theorem 10.4: SMD-WEA

Let $f : \mathbb{R}^m_{++} \times \mathbb{R}^{m\times n}_{++} \to \mathbb{R}^m_+$ be a continuous function that is homogeneous of degree 0 in (p, w) such that $p \cdot f(p) = 0$.

Then, $\forall \varepsilon \in \mathbb{R}_{++}$, there exists an economy E with n consumers with preferences $\succsim^i$ that are continuous, non-decreasing, and strictly convex and endowments $\omega \in \mathbb{R}^{m\times n}_{++}$ such that $z^N : \mathbb{R}^m_{++} \times \mathbb{R}^{m\times n}_+ \to \mathbb{R}^m$ satisfies

$$\begin{aligned} &\forall p \in \{p \in \mathbb{R}^m_{++} | \forall k \in M,\, p_k > \varepsilon\}, \\ &z^N(p, \omega) = f(p) \end{aligned}$$

[41] MWG p.611 define this property with $z(p, \omega)$. The two are equivalent in thie context.

Theorem 10.5: WEA Price

Take a *closed* set of prices P such that $P \neq \varnothing$ and $P \subseteq \mathbb{R}^m_{++}$, then there exists a pure exchange economy E with $\succsim^i$ continuous, strictly convex, and monotonic such that the Walrasian equilibrium price p^* in this economy is in P.

Consider an exchange economy E where all consumers have preferences $\succsim^i$ that are continuous, strictly convex, and strictly monotonic and for all producer $j \in \mathcal{J}$, the production possibility set Y^j satisfies constant returns to scale.

Define a group of consumers $C \subseteq N$ as a **coalition**.

Definition: $\forall C \subseteq N, \forall x \in X^n$, we say that "$C$ **blocks** $x \in X^n$" if $\exists x' \in X^n$ such that

(i) [42]$\forall i \in C$, $x'^i \succ^i x^i$

(ii) $\exists y \in Y$ such that $\sum_{i \in C} (x'^i - \omega^i) = y$ (i.e., x' is feasible.)

Definition: We say that an allocation (x^*, y^*, p^*) is **In The Core** if there is no coalition in E that blocks such allocation.

Proposition: Any Walrasian equilibrium in the economy E is in the core of E.

Consider a $k-$economy E^k such that we have for each consumer in N, there are now k identical clones of them with the same preferences $\succsim^i$ and endowment ω^i.

Proposition: An allocation (x^*, y^*, p^*) is in $Core(E^k)$ for all $k \in \mathbb{N}$ if and only if $(x^*, y^*, p^*) \in WE(E)$.

For readers who are interested in learning more about the convergence process to a WEA, consult big Kreps 14.6 and MWG 17.H, or of course, the brief discussion in this next section.

10.8 Stability of Walrasian Equilibria

Recall that by the Second Welfare Theorem, if we start out at a Pareto optimal allocation, we can achieve a WEA with wealth transfer, what does the process of shifting from the PEA

[42]Readers should note that in JR (a.k.a. Advanced Micro Economic Theory), this condition is changed to $\forall i \in C, x'^i \succsim^i x^i \wedge \exists h \in C,\, x'^h \succ^h x^h$

to WEA looks like? More generally, since most of the time our WEA is non-unique, and sometimes even unstable, if we assigned a wealth transfer, how does the allocations change, and what happens to the WEA?

Consider an economy E where we have n consumers, endowed with m goods, and j producers. For now, we will fix every thing but the endowment ω to study how equilibria move.

Take a sequence of endowment matrices $(\omega_t)_{t=1}^{\infty} \to \omega \in \mathbb{R}_+^{m\times n}$ and the corresponding sequence of economies $(E(\omega)_t)_{t=1}^{\infty}$. Denote the set of Walrasian equilibria of economy $E(\omega)_t$ as $WE(E(\omega)_t)$. Clearly, $WE(E(\omega)_t)$ is a **self-mapping** correspondence. By Kakutani's Fixed Point Theorem, if $WE(E(\omega)_t)$ is upper-hemi-continuous and $E(\omega)_w$ is compact, we would be able to find a fixed point in $WE(E(\omega)_t)$, i.e., a stable equilibrium in the economy $E(\omega)_t$. But is that necessarily true? Can we actually get *uhc* for $WE(E(\omega)_t)$? Let's discuss.

Proposition: Consider a sequence $((\omega_t,(x,y)_t,p_t))_{t=1}^{\infty} \to (\omega,(x,y),p)$ such that $\forall t \in \mathbb{N}$, $((\omega_t,(x,y)_t,p_t))_{t=1}^{\infty} \in WE(E(\omega)_t)$, then $(\omega,(x,y),p) \in WE(E(w))$. Notice that this is a bigger proposition than simply saying that $WE(E(\omega)_t)$ is compact. Recall the sequential characterization of upper-hemi-continuity, since $WE(E(\omega)_t)$ in this case is compact, it actually means that the mapping $WE(E(\omega)_t)$ is upper-hemi-continuous over ω.

By Kakutani's fixed point theorem and Blackwell's contraction conditions, we now know that if we start from some WEA of the economy E, it can dynamically adjust to an eventual equilibrium that is a fixed point in $WE(E(w))$. But what if we started out of the equilibrium, will we still get convergence to a WEA? Not necessarily.

Consider a dynamic process of price adjustment (Walras 1874; Samuelson 1947) called "**tatonnement**":

Definition: Let $\frac{dp_k}{d_t} = \lambda_k z_k^N(p,\omega)$ denote the price adjustments to excess demand where $t \in \mathbb{R}_+$ denotes continuous time and λ_k is the parameter for the speed of the adjustment.

Definition: We say that an equilibrium is **Locally Stable** if the dynamic process returns to it after any small perturbation. The equilibrium is **Unstable** otherwise.

Definition: We say that an economic system is stable if every starting point in the system converges to an equilibrium through the tatonnement process.

Remark: System stability is achieved in economies with 2 goods as well as economies with an unique WEA.

10.9 General Equilibrium under Uncertainty

Our discussion thus far has been in a competitive market where we have zero friction and perfect information/no uncertainty. What would happen to our general equilibrium results if we throw in uncertainty? Recall our discussion during choice under uncertainty. We should expect a similar consumer behavior as the risk-free case and simply study what happens when consumers consume "goods-state" as "goods".

Let the possible states of the world be denoted by a *finite set* S[43]. Consider the world with 2 time periods where the uncertainty problem needs to be solved in $t = 1$, and then state $s \in S$ is realized in $t = 2$.

Consider a separate market that provides risk alleviation during ex-ante planning so that consumers can trade for the rights to certain amount of goods given some $s \in S$ is realized. For simplicity, suppose that the market for state-contingent assets operate only at $t = 1$ and the regular goods market operate only at $t = 2$, after the state of the world have been realized.

10.9.1 Arrow-Debreu State-Contingent Securities

Definition: For each state $s \in S$ and good $k \in M$, an **Arrow-Debrue State-Contingent Security** (k, s) is an asset that entitles its owner to receive one unit of good k in the event that state s realized in $t = 2$. Denote the amount of AD securities (k, s) owned by consumer i as $r^i_{k,s} \in \mathbb{R}$.

Definition: Let $r^i \equiv (r^i_1, \ldots, r^i_{|S|}) \in \mathbb{R}^{m \times |S|}$ be consumer i's **State-Contingent Security Matrix**.

Definition: Let $\omega^i \equiv (\omega^i_1, \ldots, \omega^i_{|S|}) \in \mathbb{R}^{m \times |S|}$ denote the **State-Contingent Endowment Matrix** that consumers will have in $t = 2$.

Definition: Let $x^i \equiv (x^i_1, \ldots, x^i_{|S|}) \in \mathbb{R}^{m \times |S|}$ denote the **State-Contingent Demand Matrix** that consumers will have in $t = 2$.

[43] If you feel conflicted about this, consider the simple probability approach discussed in 7.1

Definition: Let $p \equiv (p_1, \ldots, p_{|S|}) \in \mathbb{R}^{m\times|S|}_{++}$ denote the **State-Contingent Prices** in $t = 2$.

We will assume that each consumer i in the economy have preferences over consumption across states. i.e., $\succsim^i$ provides an ordering over the set $X^{i\times|S|} = \mathbb{R}^{m\times|S|}_{+}$. For the simplicity of notations, we will denote this as X^{iS}.

Definition: The budget set for agent i given the price matrix is thus

$$B^i(p,\omega) \equiv \left\{ x^i \in X^{iS} \mid \sum_{s\in S} p_s \cdot x^i_s \leq \sum_{s\in S} p_s \cdot \omega^i_s \right\}$$

Definition: An allocation $x \in X^{|S|} \in \mathbb{R}^{m\times|S|\times n}_{+}$ is a *tensor* assigning a non-negative consumption value for each triple $(k, s, i) \in M \times S \times N$

Definition: In an Arrow-Debreu trade economy $E \equiv \left((X^{iS}, \succsim^i, \omega^i\right)_{i\in\mathcal{I}}$, an allocation $x^* \in X^{|S|}$ and prices $p^* \in \mathbb{R}^{|S|\times m}_{++}$ is an **Arrow-Debreu Equilibrium** if

(i) Consumers optimize: $\forall i \in \mathcal{I}, x^{*i} \succsim^i \hat{x}^i$, $\forall x^{*i}, \hat{x}^i \in B^i(p^*, \omega^i)$

(ii) Markets clear in every state: $\forall k \in M,\ s \in S,\ \sum_{i\in\mathcal{I}} x^{*i}_{k,s} - \sum_{i\in\mathcal{I}} \omega^i_{k,s} = \sum_{i\in\mathcal{I}} r^i_{k,s} = 0$

Remarks: The Walrasian equilibrium properties (Existence, Pareto Optimality, etc.) that we have learned apply here as well. In this case, optimality means that state-contingent commodities allow for an efficient allocation of risks.

10.9.2 Radner Model of General Equilibrium

What if we don't actually have $m \times |S|$ markets available for the ADE? Instead, suppose the state-contingent goods market is only available for 1 good (multiple goods would simply be duplicating the result). In this case, can we still get Walrasian equilibrium in $|S| + m$ markets? Yes, and the Radner model will walk us through this scenario. For simplicity, we will make good l the only good we can trade securities for.

Definition: Let $(S, N, M, \succsim, \omega)$ be an ordered pair of states, agents with preferences and endowments, and goods where the preferences are over X^{iS}. Recall that the endowment matrix realizes in $t = 2$.

Definition: Let $(x^*, r^*_l, p^*_{l,s}, q^*) \in X^{|S|} \times \mathbb{R}^{n\times|S|} \times \mathbb{R}^{|S|} \times \mathbb{R}^{|S|\times m}$ be an allocation that is an ordered tuple of state-contingent consumption, number of securities purchased, prices for the securities, and the price for the commodities. Such allocation is called a **Radner Equilibrium** if:

(i) Consumers optimize:

(i.a) The consumption bundles are feasible in every state:
$x_s^{*i} \in B^i(q_s^*, r_{l,s}^{*i}, \omega_s^i) \equiv \{x_s^i \in X^i \mid q_s^* \cdot x_s^i \leq q_s^* \cdot \omega_s^i + q_{l,s}^* r_{l,s}^i\}$

(i.b) The security purchased is feasible:
$r_l^{*i} \in B^i(p_{l,s}, \omega_{l,s}) \equiv \{r_l^i \in \mathbb{R}^{|S|} \mid \sum_{s \in S} (\omega_{l,s}^i + r_{l,s}^{*i}) \cdot p_{l,s}^* \leq \sum_{s \in S} \omega_{l,s}^i \cdot p_{l,s}^*\}$

(i.c) The allocation is the most preferred:
$x^{*i} \succsim^i \hat{x}^i$ for any $\hat{r}_l^i \in B^i(p_{l,s}, \omega_{l,s})$ and any $\hat{x}^i \in X^{iS}$ such that $\forall s \in S$, $\hat{x}_s^i \in B^i(q_s^*, \hat{r}_{l,s}^i, \omega_s^i)$

(ii) All markets clear:
$\forall s \in S$, $\sum_{i \in \mathcal{I}} r_{l,s}^{*i} = 0$ and $\forall k \in M, \forall s \in S$, $\sum_{i \in \mathcal{I}} (x_{k,s}^{*i} - \omega_{k,s}^{*i}) = 0$

Theorem 10.6: Radner Economy Efficiency

In a Radner (trade) economy in which $\forall i \in \mathcal{I}$, $X^i \in \mathbb{R}_+^{m \times |S|}$ is convex, $\omega^i \in int(X^i)$, and $\succsim^i$ is locally non-satiated, continuous, and strictly convex.
Then we know that

(i) A Radner equilibrium exists.

(ii) The Radner equilibrium allocation coincides with an Arrow-Debreu equilibrium allocation of the corresponding Arrow-Debreu economy.

Remark: This theorem gives us a lot more efficiency than the A-D framework because we can now achieve efficiency with just 1 asset market instead of a securities market for all m goods.

Remark: The efficiency is lost if markets are incomplete, i.e., if there is no goods market for certain states.

11 Social Choice Theory

Now that we have studied general equilibrium and Pareto efficiency, and briefly discussed how an economy gets there (tatonnement), we need to somewhat study how it actually happens. How does the society make the choice to be in or out of equilibrium?

We already know that aggregate demand theory fails at many places, and often times cannot reveal much about aggregate consumer behavior without restrictive assumptions. But is there more that we can say? Have we been looking at this all wrong?

Social Choice Theory is an axiomatic approach that economists have tried to study the problem of aggregating choices. Unfortunately, the results we will learn does not provide much hope. If you study *Mechanism Design* in the future, you might see something familiar where we extend these results and make the best we could.

Just like any new topic we discuss, we need to first define a list of things, so bear with me:

X	A generic set of alternatives[44]
N	The set of agents in the system/economy. $\lvert N \rvert = n \in \mathbb{N}$
P_i	One of agent i's strict preference relationship[45]
R_i	One of agent i's weak preference relationship[46]
P	A strict preference profile $P \equiv (P_1, \dots, P_n)$
R	A weak preference profile $R \equiv (R_1, \dots, R_n)$
$\mathcal{P}$	The set of all asymmetric binary relations on X
$\mathcal{R}$	The set of all complete binary relations on X
$\mathcal{P}_W$	The collection of all strict preference relations on X
$\mathcal{P}_L$	The collection of all strict preference relations on X that are also *total*[47]
$\mathcal{R}_W$	The collection of all weak preference relations on X
$\mathcal{R}_L$	The collection of all weak preference relations on X that are also *anti-symmetric*
$\mathcal{P}_W^n$	$\mathcal{P}_W \times \mathcal{P}_W \times \cdots \times \mathcal{P}_W$
$\mathcal{P}_L^n$	$\mathcal{P}_L \times \mathcal{P}_L \times \cdots \times \mathcal{P}_L$
$\mathcal{R}_W^n$	$\mathcal{R}_W \times \mathcal{R}_W \times \cdots \times \mathcal{R}_W$
$\mathcal{R}_L^n$	$\mathcal{R}_L \times \mathcal{R}_L \times \cdots \times \mathcal{R}_L$
$D_P \subseteq \mathcal{P}_W^n$	The domain of strict preference profiles
$D_R \subseteq \mathcal{R}_W^n$	The domain of weak preference profiles
$D \equiv (D_p, D_R)$	The domain of preference profiles

Definition: A **Strict Preference Aggregation Rule** is a mapping $P_{AR} : D_P \to \mathcal{P}$ where $P_{AR}(P) \in \mathcal{P}$ is the **Strict Social Preference**.[48]

Definition: A **Weak Preference Aggregation Rule** is a mapping $R_{AR} : D_R \to \mathcal{R}$ where $R_{AR}(R) \in \mathcal{R}$ is the **Weak Social Preference**.

Definition: A **Social Choice Rule** is a mapping[49]:

$$SC : D \times 2^X \setminus \{\varnothing\} \to 2^X \setminus \{\varnothing\}$$

Definition: Alternatively, a **Social Choice Function** is a function:

$$SCF : D \times 2^X \setminus \{\varnothing\} \to X$$

Definition (Simple Majority): A **Simple Majority (SM)** aggregation is $P_{SM} : \mathcal{P}_W^n \to \mathcal{P}$ defined as:

$$\forall p \in \mathcal{P}_W^n,\, \forall x, y \in X,\, xP_{SM}(P)y \iff \|\{i \in \mathcal{I} \mid xP_iy\}\| > \frac{n}{2}$$

Definition (Simple Majority Core): The **Simple Majority Core (SMC)** is the set of maximal elements $M(X, P_{SM}(P))$.

Definition (Condorcet Winner): We say that $x \in X$ is a **Condorcet winner** if

$$\forall P_{AR}(P),\, P \in \mathcal{P}_W^n,\, \forall y \in X \setminus \{x\},\, xP_{SM}(P)y$$

Additionally, we denote the Condorcet process $CW_{AR}(P)$.

[44] In JR, X is defined to be finite, but such restriction was not in mentioned in lecture. Readers are encouraged to think about and discuss the ramifications if X is not at least countably infinite.

[45] Recall that this means it is asymmetric and negatively transitive. See section 5 if you need a quick refresher.

[46] Complete and Transitive

[47] The difference between $\mathcal{P}_W$ and $\mathcal{P}_L$ is that $\mathcal{P}_L$ rules out indifference and uncomparables. The subscript L stands for linear order

[48] Notice that social preferences are not actually *preference relations* as negative transitivity may not hold.

[49] Readers should be aware that it is not uncommon to have social choice mapping domains be just D, instead of the Cartesian product seen here.

Definition (Relative Majority): A **Relative Majority** aggregation is $P_{RM} : \mathcal{P}_W^n \to \mathcal{P}$ defined as:

$$\forall p \in \mathcal{P}_W^n,\, \forall x, y \in X,\, xP_{RM}(P)y \iff \|\{i \in \mathcal{I} \mid xP_iy\}\| > \|\{i \in \mathcal{I} \mid yP_ix\}\|$$

Proposition: $\forall x, y \in X,\, xP_{SM}(P)y \;\Rightarrow\; xP_{RM}(P)y$. Also notice that

$$CW_{SM}(P) \subseteq CW_{RM}(P) \subseteq M(X, P_{RM}(P)) \subseteq M(X, P_{SM}(P))$$

Definition (Plurality): A **Plurality** aggregation is $P_{plu} : \mathcal{P}_L^n \to \mathcal{P}$ such that $\forall p \in \mathcal{P}_L^n,\, \forall x, y \in X$,

$$xP_{plu}(P)y \iff \|\{i \in \mathcal{I} \mid xP_iz,\, \forall z \in X \setminus \{x\}\}\| > \|\{i \in \mathcal{I} \mid yP_iz,\, \forall z \in X \setminus \{y\}\}\|$$

Example: Consider the following society with a preference profile over $\{a, b, c\} \subseteq X$:

$P \in \mathcal{P}_L^4$				P_{plu}	P_{SM}
$i=1$	$i=2$	$i=3$	$i=4$		
a	a	b	c	a	b
b	b	c	b	b, c	c
c	c	a	a		a

Notice that a in this case is a Condorcet loser, but it is preferred to any other alternatives by the plurality aggregation.

Desirable Axioms:

1. Independence of Irrelevant Alternatives (**IIA**): $P_{AR}(\cdot)$ satisfies **IIA** if $\forall x, y \in X,\, \forall P, P' \in D_P$ we have

$$[\{i \in \mathcal{I} \mid xP_iy\} = \{i \in \mathcal{I} \mid xP'_iy\}] \wedge [\{i \in \mathcal{I} \mid yP_ix\} = \{i \in \mathcal{I} \mid yP'_ix\}]$$
$$\Rightarrow\; xP_{AR}(P)y \iff xP_{AR}(P')y$$

2. Neutrality[50] (**N**): $P_{AR}(\cdot)$ is neutral if $\forall P, P' \in D_P,\, \forall x, y, w, z \in X$

$$[\{i \in \mathcal{I} \mid xP_iy\} = \{i \in \mathcal{I} \mid wP'_iz\}] \wedge (\{i \in \mathcal{I} \mid yP_ix\} = \{i \in \mathcal{I} \mid zP'_iw\}]$$
$$\Rightarrow\; xP_{AR}(P)y \iff wP_{AR}(P')z$$

[50] Neutrality is a strengthened version of *IIA*.

3. Monotonicity (**M**): $P_{AR}(\cdot)$ is monotonic if $\forall P, P' \in D_P$, $\forall x, y \in X$ such that

$$[\{i \in \mathcal{I} \mid xP_iy\} \subseteq \{i \in \mathcal{I} \mid xP_i'y\}] \wedge [\{i \in \mathcal{I} \mid xR_iy\} \subseteq \{i \in \mathcal{I} \mid xR_i'y\}]$$
$$\Rightarrow \; xP_{AR}(P)y \;\Rightarrow\; xP_{AR}(P')y$$

4. Positive Responsive (**PR**): $P_{AR}(\cdot)$ is positive-responsive if $\forall P, P' \in D_P$, $\forall x, y \in X$ such that

$$[\{i \in \mathcal{I} \mid xP_iy\} \subseteq \{i \in \mathcal{I} \mid xP_i'y\}] \wedge [\{i \in \mathcal{I} \mid xR_iy\} \subseteq \{i \in \mathcal{I} \mid xR_i'y\}]$$

with one of the two containments being strict/proper, then

$$xR_{AR}(P)y \;\Rightarrow\; xP_{AR}(P')y$$

5. Weak Pareto (**WP**): $P_{AR}(\cdot)$ is weak Pareto if $\forall x, y \in X$, $\forall P \in D$,

$$\forall i \in \mathcal{I},\, xP_iy \;\Rightarrow\; xP_{AR}(P)y$$

6. Decisive (**D**): $P_{AR}(\cdot)$ is decisive if $\forall P, P' \in D_P$, $\forall x, y \in X$ if

$$\{i \in \mathcal{I} \mid xP_iy\} = \{i \in \mathcal{I} \mid xP_i'y\}$$
$$\Rightarrow \; xP_{AR}(P)y \iff xP_{AR}(P')y$$

7. Anonymity (**AR**): $P_{AR}(\cdot)$ is anonymous if for every permutation of agents $\rho : N \to N$, $\forall p \in D_P$, $\forall x, y \in X, \forall i \in \mathcal{I}$, we have:

$$P_i' = P_{\rho(i)} \;\Rightarrow\; (xP_{AR}(P)y \iff xP_{AR}(P')y)$$

8. Rationality (**R**): $P_{AR}(\cdot)$ is rational if $\forall P \in D_P$, $P_{AR}(P)$ is negatively transitive.

9. No Dictator (**ND**): $P_{AR}(\cdot)$ allows for no dictators if $\forall x, y \in X$, $\forall P \in D_P$, $\nexists i \in \mathcal{I}$ such that

$$xP_iy \;\Rightarrow\; xP_{AR}(P)y$$

Theorem 11.1: May (1952)

Let $D = \mathcal{P}_W^n$, the only aggregation rule that satisfies

(i) Anonymity

(ii) Neutrality

(iii) Positive-responsive

is the simple majority aggregation.

This implies that since simple majority does not satisfy rationality, we have
A+N+PR $\Rightarrow$ ¬ R

Theorem 11.2: Arrow's Impossibility Theorem (1963)

If $|X| \geq 3$, $D_P = \mathcal{P}_W^n$, if P_{AR} satisfies **R**, **WP**, **IIA**, then there must be a dictator in the system. i.e., $\forall x, y \in X$, $\forall P \in D_P$, $\exists i \in \mathcal{I}$ such that

$$xP_iy \;\Rightarrow\; xP_{AR}(P)y$$

Proof 11.2: Arrow's Impossibility Theorem (Barbera 1980)

Let $|X| \geq 3$, $D = \mathcal{P}_W^n$, and an aggregation rule P_{AR} that satisfies **R**, **WP**, and **IIA**. We want to show that a dictator must exist for any two alternatives in X.

Step 1: Top or Bottom:
Take any $x \in X$, take $\mathcal{P}_L^{n,x} \subset \mathcal{P}_W^n$ defined as:

$$\mathcal{P}^{n,x} = \{P \in \mathcal{P}_W^n \mid \forall i \in \mathcal{I},\ \text{either } xP_iy, \forall y \in X \setminus \{x\} \text{ or } yP_ix,\ \forall y \in X \setminus \{x\}\}$$

Take some P^1 that looks like

$$\begin{matrix} x & \cdots & x & \cdot & \cdots & \cdot \\ \vdots & & \vdots & \vdots & & \vdots \\ \cdot & \cdots & \cdot & x & \cdots & x \end{matrix}$$

Lemma 1: $\forall P^1 \in \mathcal{P}^{n,x}$, $\forall y \in X \setminus \{x\}$, either $xP_{AR}(P^1)y$ or $yP_{AR}(P^1)x$

Proof: Suppose otherwise that $\exists y, z \in X \setminus \{x\}$ such that $yR_{AR}(P^1)xR_{AR}(P^1)z$. We will construct $P^2 \in \mathcal{P}^{n,x}$ such that $\forall i \in \mathcal{I}$, zP_i^2y while x stays at their original positions for every agent.

By **IIA**, we must have that $yR_{AR}(P^2)xR_{AR}(P^2)z$.

By **R**, we must then have that $yR_{AR}(P^2)z$.

BUT, by **WP**, we must also have that $zR_{AR}(P^2)y$.

So by contradiction, $\forall y \in X \setminus x$, either $xP_{AR}(P^1)y$ or $yP_{AR}(P^1)x$.

P^2 looks like:

$$\begin{array}{cccccc} x & \cdots & x & z & \cdots & z \\ z & \cdots & z & \vdots & & \vdots \\ \vdots & & \vdots & y & \cdots & y \\ y & \cdots & y & x & \cdots & x \end{array}$$

Step 2: There is a **pivot** for $x, y \in X$.
By Lemma 1, we know that $\exists P^3, P^4 \in \mathcal{P}^{n,x}$ such that

$$\forall i \in \mathcal{I},\ y \in X \setminus \{x\},\ xP_i^3y \wedge yP_i^4x$$

By **WP**, $\forall y \in X \setminus \{x\}$, $xP_{AR}(P^3)y$ and $yP_{AR}(P^4)x$. From P^3 and P^4, we can construct a sequence a profiles that slowly changes from P^3 to P^4:

$$\left(P^4, (P_1^3, P_2^4, \ldots, P_n^4), (P_1^3, P_2^3, P_3^4, \ldots, P_n^4), \ldots, (P_1^3, \ldots, P_{n-1}^3, P_n^4), P^3\right)$$

Since $yP_{AR}(P^4)x$ and $xP_{AR}(P^3)y$, we know that $\exists h \in \mathcal{I}$ such that

$$yP_{AR}((P_1^3, \ldots, P_{h-1}^3, P_h^4, \ldots, P_n^4))x$$

and

$$xP_{AR}((P_1^3, \dots, P_{h-1}^3, P_h^3, P_{h+1}^4, \dots, P_n^4))y$$

We call agent h the pivot between x and y.

Step 3: We want to show that this agent h is a dictator for y and z.

Construct a new profile $P^5 \in D_P$ such that

$$P^5 = (P_1^3, P_2^3, \dots, P_{h-1}^3, P^{\mathbf{5}}{}_h, P_{h+1}^4, \dots, P_n^4)$$

and that[a] $yP_h^5 xP_h^5 z$.

By **IIA**, we know that $yP_{AR}(P^5)x$ and since $xP_h^5 z$, we also have $xP_{AR}(P^5)z$.

By **R**, we have $yP_{AR}(P^5)z$.

By **IIA**, the ranking of $x \in X$ should not affect the aggregate ordering of y and z, so we can actually discharge the assumption of $P \in \mathcal{P}^{n,x}$ and just have $P \in D$.

Step 4: We thus know that in every aggregation rule, $\forall x, y, z \in X$, $\forall P \in D$, we can always find dictators $h, h' \in \mathcal{I}$ such that

$$(xP_h y \ \Rightarrow \ xP_{AR}(P)y) \ \wedge \ (yP_{h'} z \ \Rightarrow \ yP_{AR}(P)z)$$

But by **R**, $P_{AR}(P)$ must be transitive, meaning that we have

$$xP_h y \wedge yP_{h'} z \ \Rightarrow \ xP_{AR}(P)yP_{AR}(P)z$$

meaning that one of h, h' must actually be the dictator of x and z as well. So a dictator will always exists for any pair of alternatives (like (x, y)) as long as there is a third object (z) in X such that we can iterate this process.

$\square$

[a]Notice that this means $P^5 \notin \mathcal{P}^{n,x}$