2. Separating Hyperplane Theorems

Daisuke Oyama

Mathematical Economics

This version: June 19, 2023

Things to know from real analysis

- For $S \subset \mathbb{R}$,
  - $\alpha = \max S$ if
    - $\alpha \geq x$ for all S and
    - $\alpha \in S$.
  - $\alpha = \min S$ if
    - $\alpha \leq x$ for all S and
    - $\alpha \in S$.
  - $\alpha = \sup S$ if
    - $\alpha \geq x$ for all S and
    - if $\beta \geq x$ for all S, then $\beta \geq \alpha$.
  - $\alpha = \inf S$ if
    - $\alpha \leq x$ for all S and
    - if $\beta \leq x$ for all S, then $\beta \leq \alpha$.
  - $\sup S$ exists if S is bounded above;
    $\inf S$ exists if S is bounded below.

- Euclidean norm: for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\|x\| = \sqrt{\sum_{i=1}^n (x_i)^2}$.
- Euclidean distance: for $x, y \in \mathbb{R}^n$, $d(x, y) = \|x - y\|$.
- A sequence $\{x^k\}$ in $\mathbb{R}^n$ *converges* to $x^0 \in \mathbb{R}^n$ if for any $\varepsilon > 0$, there exists a natural number K such that $d(x^k, x^0) < \varepsilon$ for all $k \geq K$.

  In this case,

  - $\{x^k\}$ is said to be *convergent*,
  - x^0 is called the *limit* of $\{x^k\}$, and
  - we write $\lim_{k \to \infty} x^k = x^0$ or $x^k \to x^0$ as $k \to \infty$.
- A sequence $\{x^k\}$ in $\mathbb{R}^n$ is called a *Cauchy sequence* if for any $\varepsilon > 0$, there exists a natural number K such that $d(x^k, x^\ell) < \varepsilon$ for all $k, \ell \geq K$.
- A sequence $\{x^k\}$ in $\mathbb{R}^n$ is convergent if and only if it is a Cauchy sequence.

- $S \subset \mathbb{R}^n$ is *closed* if for any convergent sequence in $\{x^k\}$ in S with $x^k \to x^*$, we have $x^* \in S$.
- $S \subset \mathbb{R}^n$ is *open* if for every $x \in S$, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset S$,

  where $B_\varepsilon(x) = \{y \in \mathbb{R}^n \mid d(y, x) < \varepsilon\}$.
- $S \subset \mathbb{R}^n$ is closed if and only if $\mathbb{R}^n \setminus S$ is open.
- Basic properties:
  - $\emptyset$ and $\mathbb{R}^n$ are both closed and open.
  - The union of any family of open sets is open.
  - The intersection of a finite number of open sets is open.
  - The intersection of any family of closed sets is closed.
  - The union of a finite number of closed sets is closed.

- $x \in S$ is an *interior point* of S if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset S$.

  The set of all interior points of S is called the *interior* of S and denoted $\operatorname{int} S$.

- $\operatorname{int} S$ is the largest open set that is contained in S.

- $x \in \mathbb{R}^n$ is a *boundary point* of S if for any $\varepsilon > 0$, $B_\varepsilon(x) \cap S \neq \emptyset$ and $B_\varepsilon(x) \cap (\mathbb{R}^n \setminus S) \neq \emptyset$.

  The set of all boundary points of S is called the *boundary* of S and denoted $\operatorname{bd} S$.

- The *closure* of $S \subset \mathbb{R}^n$ is the set of all points that are the limits of convergent sequences of points in S and denoted $\operatorname{cl} S$.

- $\operatorname{cl} S$ is the smallest closed set that contains S.

- Relationships:
  - $\operatorname{cl} S = \mathbb{R}^n \setminus \operatorname{int}(\mathbb{R}^n \setminus S)$
  - $\operatorname{int} S = \mathbb{R}^n \setminus \operatorname{cl}(\mathbb{R}^n \setminus S)$
  - $\operatorname{bd} S = \operatorname{cl} S \setminus \operatorname{int} S$

- $S \subset \mathbb{R}^n$ is *bounded* if there exists $r > 0$ such that $S \subset B_r(0)$.
- $S \subset \mathbb{R}^n$ is *compact* if it is closed and bounded.
- If $S \subset \mathbb{R}$ is compact, then $\sup S \in S$ and $\inf S \in S$ (and hence $\sup S = \max S$ and $\inf S = \min S$).
- If $S \subset \mathbb{R}^n$ is bounded, then any sequence in S has a convergent subsequence.

  If in addition S is closed (i.e., S is compact), then any sequence in S has a convergent subsequence and its limit is in S.
- Conversely, if every sequence in S has a convergent subsequence whose limit is in S, then S is compact.

- A family of subsets of $\mathbb{R}^n$ is said to have the *finite intersection property* if the intersection of any finite subfamily of it is nonempty.
- For $S \subset \mathbb{R}^n$, the following conditions are equivalent:
  - S is compact.
  - For any family $(F_\lambda)_{\lambda \in \Lambda}$ of closed subsets of S that has the finite intersection property, we have $\bigcap_{\lambda \in \Lambda} F_\lambda \neq \emptyset$.

- Let $S \subset \mathbb{R}^n$, $S \neq \emptyset$.

  A function $f\colon S \to \mathbb{R}^m$ is *continuous* at $\bar{x} \in S$ if for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$d(x, \bar{x}) < \delta,\ x \in S \Rightarrow d(f(x), f(\bar{x})) < \varepsilon.$$

- $f\colon S \to \mathbb{R}^m$ is continuous on $T \subset S$ if it is continuous at every $x \in T$.
- $f\colon S \to \mathbb{R}^m$ is continuous if it is continuous on S.
- $f\colon S \to \mathbb{R}^m$ is continuous at $\bar{x} \in S$ if and only if for any sequence $\{x^k\}$ in S such that $x^k \to \bar{x}$, we have $f(x^k) \to f(\bar{x})$.
- Examples:
  - $(x, y) \mapsto d(x, y)$ is continuous.
  - For $A \in \mathbb{R}^{m \times n}$, $x \mapsto Ax$ is continuous.

Proposition 2.1 (Weierstrass' Theorem)

Suppose that $S \subset \mathbb{R}^n$, $S \neq \emptyset$, *is compact and* $f\colon S \to \mathbb{R}$ *is continuous. Then* $\max_{x\in S} f(x)$ *and* $\min_{x\in S} f(x)$ *exist, i.e., there exist* $x^*, x^{**} \in S$ *such that* $f(x^{**}) \leq f(x) \leq f(x^*)$ *for all* $x \in S$.

Convex Sets

Definition 2.1

$C \subset \mathbb{R}^n$ is *convex* if for any $x, y \in C$ and $\lambda \in [0, 1]$, we have $\lambda x + (1 - \lambda) y \in C$.

Properties of Convex Sets

Proposition 2.2

Suppose that $C, D \subset \mathbb{R}^n$ *are convex.*

- $C + D = \{x + y \mid x \in C,\ y \in D\}$ *is convex.*
- *For* $\alpha \in \mathbb{R}$, $\alpha C = \{\alpha x \mid x \in C\}$ *is convex.*

Proposition 2.3

The intersection of any family of convex sets is convex.

Proposition 2.4

If $C \subset \mathbb{R}^n$ is convex, then $\operatorname{cl} C$ is also convex.

Proof

- $B_\varepsilon(0) = \{y \in \mathbb{R}^n \mid \|y\| < \varepsilon\}$ is convex.
- Then $\operatorname{cl} C = \bigcap_{\varepsilon > 0}(C + B_\varepsilon(0))$ is convex if C is convex.

Proposition 2.5

If $C \subset \mathbb{R}^n$ is convex, then $\operatorname{int} C$ is also convex.

- For $S \subset \mathbb{R}^n$, the set of all convex combinations of finite subsets of S is called the *convex hull* of S and denoted by $\mathrm{conv}(S)$.
- $\mathrm{conv}(S)$ is the smallest convex set that contains S.

Carathéodory's Theorem

Proposition 2.6 (Carathéodory's Theorem)

1. *For* $S \subset \mathbb{R}^n$, $S \neq \{0\}$, *each* $x \in \mathrm{cone}(S)$ *is written as a conic combination of linear independent elements of* S.

2. *For* $S \subset \mathbb{R}^n$, *each* $x \in \mathrm{conv}(S)$ *is written as a convex combination of at most* $n + 1$ *elements of* S.

Proof of Part 1

- Immediate from Proposition 1.6 (Fundamental Theorem of Linear Inequalities).

Proof of Part 2

- Let $x \in \operatorname{conv}(S)$.

  Then we have $x = \sum_{j=1}^{J} \lambda_j x^j$ for some $x^1, \ldots, x^J \in S$ and $\lambda_1, \ldots, \lambda_J \geq 0$, $\sum_{j=1}^{J} \lambda_j = 1$.

- Consider $T = \{(x^1, 1), \ldots, (x^J, 1)\} \subset \mathbb{R}^{n+1}$.

  Then $(x, 1) \in \operatorname{cone}(T)$.

- By part 1, there is an LI subset $T' \subset T$ such that $(x, 1) = \sum_{j \in T'} \mu_j (x^j, 1)$ with $\mu_j \geq 0$, where $|T'| \leq n + 1$.

- From the 1st through nth coordinates we have $x = \sum_{j \in T'} \mu_j x^j$, while from the $(n+1)$st coordinate we have $\sum_{j \in T'} \mu_j = 1$.

Convex Hull of a Compact Set

Proposition 2.7

If $S \subset \mathbb{R}^n$ *is bounded, then* $\mathrm{cl}(\mathrm{conv}(S)) = \mathrm{conv}(\mathrm{cl}(S))$.
In particular, if S *is compact, then* $\mathrm{conv}(S)$ *is compact.*

Proof

- Since $\text{conv}(S) \supset S$, we have $\text{cl}(\text{conv}(S)) \supset \text{cl}(S)$.

  Since $\text{cl}(\text{conv}(S))$ is convex (Proposition 2.4), we have $\text{cl}(\text{conv}(S)) \supset \text{conv}(\text{cl}(S))$.

- Since $S \subset \text{cl}(S)$, we have $\text{conv}(S) \subset \text{conv}(\text{cl}(S))$.

  We want to show that $\text{conv}(\text{cl}(S))$ is closed if S is bounded.

- Let $\{x^k\} \subset \text{conv}(\text{cl}(S))$, and assume $x^k \to \bar{x}$.

- By Carathéodory's Theorem (Proposition 2.6 part 2), each x^k is written as

$$x^k = \alpha_1^k x^{k,1} + \cdots + \alpha_{n+1}^k x^{k,n+1},$$

  where

  - $(\alpha_1^k, \ldots, \alpha_{n+1}^k) \in \Delta = \{\alpha \in \mathbb{R}^{n+1} \mid \alpha_i \geq 0,\ \sum_i \alpha_i = 1\}$,
  - $x^{k,1}, \ldots, x^{k,n+1} \in \text{cl}(S)$.

- Since Δ and $\mathrm{cl}(S)$ are compact, there exists a sequence $\{k(\ell)\}$ such that the limits $\bar{\alpha}_i = \lim_{\ell \to \infty} \alpha_i^{k(\ell)}$ and $\bar{x}^i = \lim_{\ell \to \infty} x^{k(\ell),i}$ exist where $(\bar{\alpha}_1, \ldots, \bar{\alpha}_{n+1}) \in \Delta$ and $\bar{x}^1, \ldots, \bar{x}^{n+1} \in \mathrm{cl}(S)$.
- Hence,

$$\bar{x} = \bar{\alpha}_1 \bar{x}^1 + \cdots + \bar{\alpha}_{n+1} \bar{x}^{n+1},$$

  so that $\bar{x} \in \mathrm{conv}(\mathrm{cl}(S))$.

Separating Hyperplane Theorems

- The textbook proves the strict separating hyperplane theorem from scratch.
- It then states the weak separating hyperplane theorem without proof, saying "The proof is similar to the previous one."

  (In fact, the proof is far from "similar".)
- Here, we prove the weak separating hyperplane theorem by Farkas' Lemma (which we proved by an algebraic argument).
- Then we prove the strict version from the weak version.

Weak Separating Hyperplane Theorem

Proposition 2.8 (Weak Separating Hyperplane Theorem)

Suppose that $C \subset \mathbb{R}^n$ is a convex set, and that $b \notin C$. Then there exists $h \in \mathbb{R}^n$, $h \neq 0$ such that

$hx \leq hb$ *for all* $x \in C$.

- The proof below is an adoption of a proof in some lecture notes by Atsushi Kajii (which proves this theorem from the strict version).
- A similar argument (similar to Kajii's) is also found in Berkovitz, *Convexity and Optimization in* $\mathbb{R}^n$, Chapter II, Theorem 3.2.

Proof

- Write $P^0 = \{h \in \mathbb{R}^n \mid \|h\| = 1\}$, which is compact.
- Let $C \subset \mathbb{R}^n$ be convex and $b \notin C$.
- For each $x \in C$, let

  $$P_x = \{h \in P^0 \mid hx \leq hb\},$$

  which is a closed subset of P^0.

  We want to show that $\bigcap_{x \in C} P_x \neq \emptyset$.
- We show that the family $\{P_x\}_{x \in C}$ of closed subsets of compact set P^0 has the finite intersection property.

  Take any $x^1, \ldots, x^m \in C$.

  Write $A = \left[x^1 \cdots x^m\right] \in \mathbb{R}^{n \times m}$.

- Since $b \notin \mathrm{conv}(A)$ ($\subset C$), there exists no $\alpha \in \mathbb{R}^m$ such that $b = A\alpha$, $\mathbf{1}\alpha = 1$, and $\alpha \geq 0$ (where $\mathbf{1} \in \mathbb{R}^m$ is the vector of ones), or such that $\begin{bmatrix} b \\ 1 \end{bmatrix} = \begin{bmatrix} A \\ \mathbf{1}^{\mathrm{T}} \end{bmatrix} \alpha$ and $\alpha \geq 0$.
- Then by Farkas' Lemma, there exist $h \in \mathbb{R}^n$ and $k \in \mathbb{R}$ such that $\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} A \\ \mathbf{1}^{\mathrm{T}} \end{bmatrix} \leq 0$ and $\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} b \\ 1 \end{bmatrix} > 0$, or

  $hx^j \leq -k < hb$ for all $j = 1, \ldots, m$,

  so that $h \in \bigcap_{j=1}^m P_{x^j}$.
- Thus, $\bigcap_{j=1}^m P_{x^j} \neq \emptyset$.
- Hence, by the compactness of P^0, we have $\bigcap_{x \in C} P_x \neq \emptyset$, as desired.

Proposition 2.9

Suppose $C, D \subset \mathbb{R}^n$, $C, D \neq \emptyset$, *are convex, and that* $C \cap D = \emptyset$. *Then there exists* $h \in \mathbb{R}^n$, $h \neq 0$ *such that*

$$hx \leq hy \text{ for all } x \in C \text{ and } y \in D.$$

Proof

- Let $K = C - D$ $(= \{x - y \mid x \in C,\ y \in D\})$. Then
  - $K \neq \emptyset$ ($\because C, D \neq \emptyset$);
  - K is convex ($\because C$ and D convex);
  - $0 \notin K$ ($\because C \cap D = \emptyset$).
- Therefore, by the weak separating hyperplane theorem, there exists $h \in \mathbb{R}^n$, $h \neq 0$, such that

$$hz \leq h0 \text{ for all } z \in K,$$

or

$$hx \leq hy \text{ for all } x \in C \text{ and } y \in D.$$

Strict Separating Hyperplane Theorem

Proposition 2.10 (Strict Separating Hyperplane Theorem)

Suppose that $C \subset \mathbb{R}^n$ is a closed convex set, and that $b \notin C$. Then there exist $h \in \mathbb{R}^n$, $h \neq 0$, and $\beta \in \mathbb{R}$ such that

$$hx \leq \beta < hb \text{ for all } x \in C.$$

Proof

- Let $b \notin C$.
- By the closedness of C, there exists $\bar{\varepsilon} > 0$ such that $C \cap B_{\bar{\varepsilon}}(b) = \emptyset$.
- By the convexity of C (and $B_{\bar{\varepsilon}}(b)$), it follows from Proposition 2.9 that there exists $h \in \mathbb{R}^n$, $h \neq 0$ such that $hx \leq hy$ for all C and all $y \in B_{\bar{\varepsilon}}(b)$.

  Normalize h so that $\|h\| = 1$.
- Letting $y = b - \frac{\bar{\varepsilon}}{2}h$, we have $hx \leq hb - \frac{\bar{\varepsilon}}{2}$ for all $x \in C$, where $hb - \frac{\bar{\varepsilon}}{2} < hb$.
- Finally, let $\beta = hb - \frac{\bar{\varepsilon}}{2}$.

Proposition 2.11

Suppose $C, D \subset \mathbb{R}^n$, $C, D \neq \emptyset$, *are convex and closed, and that* $C \cap D = \emptyset$.
If C *or* D *is bounded, then there exist* $h \in \mathbb{R}^n$, $h \neq 0$, *and* $\beta \in \mathbb{R}$ *such that*

$$hx < \beta < hy \text{ for all } x \in C \text{ and } y \in D.$$

- The boundedness of C or D is indispensable.

Proof

- Let $K = C - D$ $(= \{x - y \mid x \in C,\ y \in D\})$. Then
  - $K \neq \emptyset$ $(\because C, D \neq \emptyset)$;
  - K is convex $(\because C$ and D convex$)$;
  - $0 \notin K$ $(\because C \cap D = \emptyset)$.
- Suppose that C is bounded and hence is compact by the closedness of C.

  We want to show that K is closed.

- Take any sequence $\{z^k\}$ in K, and assume that $z^k \to z^*$.
  - For each k, let $x^k \in C$ and $y^k \in D$ be such that $z^k = x^k - y^k$.
  - By the compactness of C, there are a subsequence of $\{x^k\}$ (again denoted $\{x^k\}$) and $x^* \in C$ such that $x^k \to x^*$.
  - Then $y^k = x^k - z^k$ converges to some y^*, where $y^* \in D$ by the closedness of D.
  - Then we have $z^k = x^k - y^k \to z^* = x^* - y^*$, and hence $z^* \in K$.
  - This proves that K is closed.

- Therefore, by the strict separating hyperplane theorem, there exist $h \in \mathbb{R}^n$, $h \neq 0$, and $\beta' \in \mathbb{R}$ such that

  $$hz < \beta' < h0 \text{ for all } z \in K,$$

  or

  $$hx < hy + \beta' < hy \text{ for all } x \in C \text{ and } y \in D.$$

- Then let, for example, $\beta = \inf_{y \in D} hy + \frac{\beta'}{2}$.

Extreme Points and Extreme Rays

Definition 2.2

For $S \subset \mathbb{R}^n$, $x \in S$ is an *extreme point* of S if

$$x = \lambda y + (1-\lambda)z,\ \ y, z \in S,\ \ \lambda \in (0,1) \implies y = z = x.$$

Definition 2.3

For $S \subset \mathbb{R}^n$,
$r \in S$, $r \neq 0$, is a *ray* of S if $x + \lambda r \in S$ for all $x \in S$ and $\lambda \geq 0$;
$r \in S$ is an *extreme ray* of S if

$$r = \lambda u + (1-\lambda)v,\ \ u, v\text{: rays of } S,\ \ \lambda \in (0,1)$$
$$\implies u = \alpha v \text{ for some } \alpha > 0.$$

Krein-Milman Theorem

- Denote the set of extreme points of C by $\operatorname{ext}(C)$.

Proposition 2.12 (Krein-Milman Theorem)

Let $C \subset \mathbb{R}^n$, $C \neq \emptyset$, be a compact convex set.
Then $C = \operatorname{conv}(\operatorname{ext}(C))$.

- In the proof given in the textbook, I could not prove the closedness of K from the induction hypothesis (rather than proving the Krein-Milman Theorem itself by a different proof).

- We prove the theorem in a stronger form.

Proposition 2.13 (Krein-Milman Theorem)

Let $C \subset \mathbb{R}^n$, $C \neq \emptyset$, be a compact convex set.
Then each $x \in C$ is written as a convex combination of at most $n+1$ extreme points of C.

- The proof is by induction on the dimension of C.

Dimension of a set

Definition 2.4

$\{x^0, \ldots, x^m\} \subset \mathbb{R}^n$ is *affinely independent* if $\{x^1 - x^0, \ldots, x^m - x^0\}$ is LI.

Definition 2.5

For $S \subset \mathbb{R}^n$, the *dimension* of S, $\dim S$, is the largest number m for which S contains some affinely independent vectors $x^0, \ldots, x^m$.

- For any $x^0 \in S$, $\dim(S) = \operatorname{rank}(S - \{x^0\})$.
- $\dim \mathbb{R}^n = n$ (take 0 and the unit vectors $e^1, \ldots, e^n$).
- For a hyperplane $H \subset \mathbb{R}^n$, i.e., $H = \{x \in \mathbb{R}^n \mid hx = \beta\}$ for some $h \in \mathbb{R}^n$, $h \neq 0$, and $\beta \in \mathbb{R}$, $\dim H = n - 1$.
  - For any $x^0 \in H$, let $H^0 = H - \{x^0\} = \{x \in \mathbb{R}^n \mid hx = 0\}$.
  - $H^0 = \ker(h^{\mathrm{T}})$ and $\operatorname{rank}(\operatorname{span}(h)) = 1$, and hence $\dim(H) = \operatorname{rank}(H^0) = n - 1$.

Proof of Proposition 2.13

- We prove by induction on the dimension of C.
- If $\dim(C) = 0$, where C is a singleton set, the statement is obviously true.
- Assume that the statement is true for any compact convex set C with $\dim(C) \leq m - 1$.
- Suppose that $\dim(C) = m$. Denote $K = \operatorname{conv}(\operatorname{ext}(C))$.

  We can embed C into $\mathbb{R}^m$, so that we can assume $C \subset \mathbb{R}^m$ (where the structure of convex combinations does not change).

Claim 1

Each $x \in \operatorname{bd} C$ is written as a convex combination of at most m extreme points of C.

Proof

- Take any $\bar{x} \in \operatorname{bd} C$.

  By the weak separating hyperplane theorem (applied to $\operatorname{int} C$ which is convex), there exists a hyperplane $H = \{x \in \mathbb{R}^m \mid hx = h\bar{x}\}$ such that $hy \leq h\bar{x}$ for all $y \in C$.

- Since $C \cap H$ is compact and convex and $\dim(C \cap H) \leq m - 1$, by the induction hypothesis there are m extreme points $y^1, \ldots, y^m$ of $C \cap H$ such that $\bar{x} \in \operatorname{conv}(\{y^1, \ldots, y^m\})$.

- We want to show that $y^1, \ldots, y^m$ are extreme points of C.

- Let $y^i = \lambda z + (1-\lambda)w$, $z, w \in C$, and $\lambda \in (0,1)$.
- Then we have

$$\begin{aligned} h\bar{x} = hy^i &= \lambda hz + (1-\lambda)hw \\ &\leq \lambda h\bar{x} + (1-\lambda)h\bar{x} = h\bar{x}. \end{aligned}$$

  Thus, the inequality in fact holds with equality, and hence, $hz = hw = h\bar{x}$, meaning that $z, w \in H$.
- Since y^i is an extreme point of $C \cap H$, it must be that $z = w = y^i$.

  This proves that each y^i is an extreme point of C.

  [End of the proof of Claim 1]

- Claim 1 in particular implies that $\mathrm{ext}(C) \neq \emptyset$.

  Fix any $x^0 \in \mathrm{ext}(C)$.

- Take any $x \in C$.

  If $x = x^0$, we are done, so assume that $x \neq x^0$.

- Let $\alpha_0 = \max\{\alpha \in \mathbb{R} \mid x^0 + \alpha(x - x^0) \in C\} \geq 1$, which is well defined by the compactness of C.

  Then $y = x^0 + \alpha_0(x - x^0) \in \mathrm{bd}\, C$.

- Then by Claim 1, there exist $x^1, \ldots, x^m \in \mathrm{ext}(C)$ such that $y = \sum_{i=1}^m \alpha_i x^i$ for some $\alpha_1, \ldots, \alpha_m \geq 0$ with $\sum_{i=1}^m \alpha_i = 1$.

- Then we have

$$
\begin{aligned}
x &= x^0 + \frac{1}{\alpha_0}(y - x^0) \\
&= \frac{\alpha_0 - 1}{\alpha_0} x^0 + \frac{1}{\alpha^0} y = \frac{\alpha_0 - 1}{\alpha_0} x^0 + \frac{1}{\alpha^0} \textstyle\sum_{i=1}^m \alpha_i x^i,
\end{aligned}
$$

  where $\frac{\alpha_0 - 1}{\alpha_0}, \frac{1}{\alpha^0}\alpha_i \geq 0$ and $\frac{\alpha_0 - 1}{\alpha_0} + \frac{1}{\alpha^0} \sum_{i=1}^m \alpha_i = 1$, as desired.

Application: Walrasian Equilibrium in Exchange Economies

- Goods $1, \ldots, n$
- Agents $A = \{1, \ldots, m\}$
- For each agent $i \in A$:
  - Endowment $w^i \in \mathbb{R}^n_+$

    Assume $w^i \gg 0$.
  - Utility function $U^i \colon \mathbb{R}^n_+ \to \mathbb{R}$

    Assumed to be
    - continuous;
    - strictly quasi-concave; and
    - strictly increasing: i.e., if $y \geq x$ and $y \neq x$, then $U^i(y) > U^i(x)$.
- Let $M \in \mathbb{R}^n_+$ be such that $M \geq \sum_{i \in A} w^i$.
  (In particular, $M \gg w^i$ for all $i \in A$.)
- $p \in \mathbb{R}^n_+$: Price vector (to be determined in equilibrium)

- Demand function of agent i:

$$d^i(p) = \arg\max\{U^i(x) \mid x \in \mathbb{R}^n_+,\ px \le pw^i,\ x \le M\}$$

  - "$x \le M$" is a non-standard constraint, which makes the domain compact even when the prices of some goods are zero.
  - By the continuity of U^i, the right-hand side is nonempty.
  - By the strict quasi-concavity of U^i, the right-hand side is a singleton set.

    $\rightarrow$ We regard $d^i(p)$ as a function (instead of correspondence).

Observation 1

For any $p \in \mathbb{R}^n_+$, if $px < pw^i$ and $x \leq M$, then $U^i(x) < U^i(d^i(p))$; in particular, $pd^i(p) = pw^i$.

- This holds if U^i satisfies local insatiability within $\{x \in \mathbb{R}^n_+ \mid x \leq M\}$ (denote this set by M),

  i.e., the property that for any $x \in M$ and any $\varepsilon > 0$, there exists $x' \in M$ such that $\|x' - x\| < \varepsilon$ and $U^i(x') > U^i(x)$.

  (Local insatiability within $\mathbb{R}^n_+$ is not sufficient.)

Proof

- Let $p \neq 0$.
- Let $x \in \mathbb{R}^n_+$ be such that $px < pw^i$ and $x \leq M$.

  We want to show that such an x is not optimal.
- By $M \gg w^i$, there must exist j such that $p_j > 0$ and $x_j < M_j$ (otherwise we would have $px > pw^i$).
- Let $x' \in \mathbb{R}^n_+$ be such that x'_j is slightly larger than x_j (while $x'_k = x_k$ for all $k \neq j$) so that we still have $px' \leq pw^i$ and $x'_j \leq M_j$.
- By strict monotonicity of U^i, we have $U^i(x') > U^i(x)$.

  This means that x is not optimal.

Observation 2
$d^i(tp) = d^i(p)$ for any $t > 0$.

Observation 3
If $U^i(x) > U^i(d^i(p))$ and $x \leq M$, then $px > pw^i$.

- We normalize a price vector $p \geq 0$, $p \neq 0$, so that $\sum_j p_j = 1$, or consider p as an element of $\Delta = \{p \in \mathbb{R}^n_+ \mid \sum_j p_j = 1\}$.

Definition 2.6

A pair of price vector $p \in \mathbb{R}^n_+$ and allocation $X = (x^1, \ldots, x^m) \in (\mathbb{R}^n_+)^m$ is a *Walrasian equilibrium* if

- [utility maximization]
  $x^i = d^i(p)$ for all $i \in A$, and
- [market clearing]
  $\sum_{i \in A} x^i = \sum_{i \in A} w^i$.

- The market clearing condition should be imposed as an inequality (i.e., $\sum_{i \in A} x^i \leq \sum_{i \in A} w^i$) if we do not assume monotonicity of U^i.

Lemma 2.14

$d^i(p)$ is continuous on Δ.

- By the continuity of U^i and the "continuity" of the constraint correspondence $p \mapsto \{x \in \mathbb{R}^n_+ \mid px \leq pw^i,\ x \leq M\}$.

Proof

- Let $\{p^k\}$ be a sequence in Δ and assume that $p^k \to p^* \in \Delta$.
- Write $x^k = d^i(p^k)$.

  Since it is contained in the compact set $\{x \in \mathbb{R}^n_+ \mid x \leq M\}$, we assume that $\{x^k\}$ is convergent with limit $x^* \in \mathbb{R}^n_+$, $x^* \leq M$.

  We want to show that $d^i(p^*) = x^*$.
- First, since $p^k x^k \leq p^k w^i$ for all k, by $k \to \infty$ we have $p^* x^* \leq p^* w^i$.

- Second take any $z \in \mathbb{R}^n_+$ such that $p^* z \leq p^* w^i$ and $z \leq M$.

  We want to show that $U^i(z) \leq U^i(x^*)$.

- For any $\varepsilon > 0$, let $z^\varepsilon \in \mathbb{R}^n_+$ be such that $\|z^\varepsilon - z\| < \varepsilon$, $p^* z^\varepsilon < p^* w^i$, and $z^\varepsilon \leq M$.

  (Note that $p^* w^i > 0$ since $w^i \gg 0$ by assumption.)

- Let K be such that $p^k z^\varepsilon \leq p^k w^i$ for all $k \geq K$.

- Then by optimality we have $U^i(z^\varepsilon) \leq U^i(x^k)$.

- Letting $k \to \infty$, we have $U^i(z^\varepsilon) \leq U^i(x^*)$ by continuity of U^i.

- Finally, letting $\varepsilon \to 0$, we have $U^i(z) \leq U^i(x^*)$ again by continuity of U^i.

- Define the function $E(p) = \sum_{i \in A} d^i(p) - \sum_{i \in A} w^i$.

  $\cdots$ Excess demand function

  Continuous by Lemma 2.14.

Lemma 2.15 (Walras' Law)
For any $p \in \mathbb{R}^n_+$, $pE(p) = 0$.

- By Observation 1.

Existence of Walrasian equilibrium

Proposition 2.16

There exists a Walrasian equilibrium.

- $p \in \Delta$ is a Walrasian equilibrium price vector if and only if $E(p) = 0$, or it is a fixed point of the function $p + E(p)$.
- But $p + E(p) \notin \Delta$ in general.

  We will modify this function so that the value is in Δ.
- Then use Brouwer's Fixed Point Theorem.

Brouwer's Fixed Point Theorem

Proposition 2.17

Suppose that $X \subset \mathbb{R}^N$ *is a nonempty, compact, and convex set, and that* $f\colon X \to X$ *is a continuous function from* X *into itself. Then* f *has a fixed point, i.e., there exists* $x \in X$ *such that* $x = f(x)$.

Proof of Proposition 2.16

- Write $E_j^+(p) = \max\{E_j(p), 0\}$, which is continuous in p.
- Define the function $f\colon \Delta \to \Delta$ by

$$f_j(p) = \frac{p_j + E_j^+(p)}{1 + \sum_{j=1}^m E_j^+(p)},$$

  which is continuous, mapping the compact set Δ to itself.
- By Brouwer's Fixed Point Theorem, f has a fixed point $p \in \Delta$:

$$p_j = \frac{p_j + E_j^+(p)}{1 + \sum_{j=1}^m E_j^+(p)}.$$

- Then by Walras' Law $pE(p) = 0$, we have

$$0 = \sum_j p_j E_j(p) = \frac{\sum_j p_j E_j(p) + \sum_j E_j^+(p) E_j(p)}{1 + \sum_j E^+ j(p)}$$
$$= \frac{\sum_j E_j^+(p) E_j(p)}{1 + \sum_j E^+ j(p)},$$

and therefore $\sum_j E_j^+(p) E_j(p) = 0$.

- Since

$$E_j^+(p) E_j(p) = \begin{cases} E_j(p)^2 & \text{if } E_j(p) > 0, \\ 0 & \text{if } E_j(p) \leq 0, \end{cases}$$

it must be that $E_j(p) \leq 0$ for all j.

- Finally, we want to show that $E_j(p) = 0$ for all j (by strict monotonicity of U^i).
- By Walras' Law, $\sum_j p_j E_j(p) = 0$, where $p_j E_j(p) \leq 0$ as shown.
- If $E_j(p) < 0$, then $p_j = 0$,

  but by monotonicity of U^i, we would have $d_j^i(p) = M_j \geq \sum_{h \in A} w_j^h$ (> 0) for all $i \in A$, which violates $E_j(p) \leq 0$.

Pareto Efficiency of Walrasian equilibrium

- An allocation $X = (x^1, \ldots, x^m) \in (\mathbb{R}^n_+)^m$ is *feasible* if $\sum_{i \in A} x^i \leq \sum_{i \in A} w^i$.
- An allocation Y *Pareto dominates* an allocation X if
  - $U^i(y^i) \geq U^i(x^i)$ for all $i \in A$, and
  - $U^i(y^i) > U^i(x^i)$ for some $i \in A$.
- A feasible allocation X is *Pareto efficient* (or Pareto optimal) if there exists no feasible allocation Y that Pareto dominates X.

First Fundamental Theorem of Welfare Economics

Proposition 2.18

If (p, X) is a Walrasian equilibrium, then X is Pareto efficient.

- Uses only Observation 1.

Proof

- Suppose that an allocation Y Pareto dominates X, i.e.,

$$U^i(y^i) \geq U^i(x^i) \text{ for all } i \in A, \tag{1}$$
$$U^i(y^i) > U^i(x^i) \text{ for some } i \in A. \tag{2}$$

  We want to show that Y is not feasible.

- If $y^i \not\leq M$ for some $i \in A$, then clearly Y is not feasible.

  Suppose that $y^i \leq M$ for all $i \in A$.

- By (1) and Observation 1, we have

$$py^i \geq pw^i \text{ for all } i \in A.$$

- By (2), we have

$$py^i > pw^i \text{ for some } i \in A.$$

- Therefore, we have

$$p\left(\sum_{i\in A} y^i - \sum_{i\in A} w^i\right) = \sum_{i\in A}(py^i - pw^i) > 0.$$

- This implies that $\sum_{i\in A} y^i \leq \sum_{i\in A} w^i$ does not hold, i.e., Y is not feasible,

  for, we would have $p\left(\sum_{i\in A} y^i - \sum_{i\in A} w^i\right) \leq 0$ otherwise.

Second Fundamental Theorem of Welfare Economics

Proposition 2.19

Suppose that $X = (w^1, \ldots, w^m)$ *is Pareto efficient. Then there exists* $p \in \mathbb{R}^n_+$ *such that* (p, X) *is a Walrasian equilibrium.*

- Uses
  - quasi-concavity,
  - local insatiability, and
  - continuity of U^i; and
  - $w^i \gg 0$.

Proof

- Define

$$\hat{S}^i = \{y^i \in \mathbb{R}^n_+ \mid U^i(y^i) > U^i(w^i)\},$$

  and define $\hat{S} = \sum_{i \in A} \hat{S}^i$, which is a convex set by the quasi-concavity of U^i's.

- By the Pareto efficiency of $X = (w^1, \ldots, w^m)$, $\hat{S} \cap (\{\sum_{i \in A} w^i\} - \mathbb{R}^n_+) = \emptyset$.

- By the weak separating hyperplane theorem, there exists $p \in \mathbb{R}^n$, $p \neq 0$, such that

$$py \geq p\left(\textstyle\sum_{i \in A} w^i - z\right) \text{ for all } y \in \hat{S} \text{ and } z \geq 0.$$

- Since this holds for all $z \geq 0$, it must be that $p \geq 0$.

- We want to show that (p, X) is a Walrasian equilibrium.

- Fix any $i \in A$.
  Suppose that $y^i \in \mathbb{R}^n_+$, $U^i(y^i) > U^i(w^i)$, and $y^i \leq M$.
- For each $j \neq i$, by strict monotonicity of U^j (local insatiability is sufficient) we have y^j arbitrarily close to w^j such that $U^j(y^j) > U^i(w^i)$.
- Then $\sum_j y^j \in \hat{S}$, and therefore,
  $p(y^i + \sum_{j \neq i} y^j) \geq p(w^i + \sum_{j \neq i} w^j)$.
- Letting $y^j \to w^j$ for all $j \neq i$, we have $py^i \geq pw^i$.

  (We have shown that (p, X) is a "quasi-equilibrium".)
- We want to show that if $y^i \in \mathbb{R}^n_+$, $U^i(y^i) > U^i(w^i)$, and $y^i \leq M$, then $py^i > pw^i$.

- Suppose that $y^i \in \mathbb{R}^n_+$, $U^i(y^i) > U^i(w^i)$, and $y^i \leq M$.
- By the continuity of U^i, $U^i(\alpha y^i) > U^i(w^i)$ for some $\alpha < 1$.

  Then, as we have shown, we must have $p(\alpha y^i) \geq pw^i$.
- Since $w^i \gg 0$ and $p \geq 0$, $p \neq 0$, we have $0 < pw^i \leq \alpha(py^i) < py^i$.