Economic Theory and Some Useful Math

Author: Wenxiao Yang

Institute: Haas School of Business, University of California Berkeley

Date: 2023

All models are wrong, but some are useful.

Contents

Chapter 1 Economic Theory Foundation

Based on

- Mas-Colell, Whinston, and Green, Microeconomic Theory, Oxford University Press (1995).
- UIUC ECON 530 21Fall, Nolan H. Miller
- UC Berkeley ECON 201A 23Fall
- UC Berkeley MATH 272 23Fall, Alexander Teytelboym
- Jehle, G., Reny, P.: Advanced Microeconomic Theory . Pearson, 3rd ed. (2011). Ch. 6.
- Notes on Social Choice and Welfare, Alejandro Saporiti
- Yu, N. N. (2012). A one-shot proof of Arrow's impossibility theorem. *Economic Theory*, 523-525.

1.1 Preferences and Utility Functions

1.1.1 Rational Preference

Definition 1.1 (Complete Preference)

Completeness: $\succeq$ is complete iff $\forall x, y \in X$, $x \succeq y$ or $y \succeq x$. ♣

The completeness means

- Any two bundles can be compared
- Indifference is allowed

Definition 1.2 (Transitive Preference)

Transitivity: $\succeq$ is transitive iff $\forall x, y, z \in X$, if $x \succeq y$ and $y \succeq z$, then $x \succeq z$ ♣

The transitivity

- like transitivity of the real numbers
- extends pairwise preferences to longer chains in the logical way.

Definition 1.3 (Rantional Preference)

Rationality: $\succeq$ is **rational** if and only if it is **complete** and **transitive**. ♣

1.1.2 Utility Function $\Leftrightarrow$ Rational Preference

Definition 1.4 (Unitility Function)

We can say a function $u : X \to \mathbb{R}$ represents $\succeq$ if $\forall x, y \in X$,

$$x \succeq y \Leftrightarrow u(x) \geq u(y)$$

Proposition 1.1 (Rational $\succeq \Rightarrow \exists u(\cdot)$)

If $\exists$ a function $u : X \to \mathbb{R}$ represents $\succeq$, then $\succeq$ is rational (i.e., completeness and transitivity)

Note *The reverse may not true.*

Let $\mathcal{B} = 2^X$ (all subsets of X) and $B \in \mathcal{B}$ be the all potential alternatives that can be chosen.

The choice of an agent can be represented by $C(B) \subseteq B, \forall B \in \mathcal{B}$.

$\succsim$		u
monotone	$\Longrightarrow$	nondecreasing
strongly monotone	$\Longrightarrow$	strictly increasing
continuous	$\Longrightarrow$	continuous (Debreu's Theorem)
convex	$\Longrightarrow$	quasi-concave (but not concave)
strictly convex	$\Longrightarrow$	strictly concave (and strictly quasi-concave)
homothetic (and continuous)	$\Longrightarrow$	continuous and homogeneous
(so-called) quasi-linear	$\Longrightarrow$	quasi-linear
(so-called) differentiable	$\Longrightarrow$	differentiable
separable	$\Longrightarrow$	separable (form)
strongly separable	$\Longrightarrow$	additively separable (form)

Figure 1.1: Properties of Preference and Utility Function

1.1.3 Monotone Preference

Definition 1.5 (Monotone $\succeq$)

$\succeq$ is **monotone** if $x, y \in X$ with $x \geq y \Rightarrow x \succeq y$ (and $x > y \Rightarrow x \succ y$).

Proposition 1.2 (Monotone $\succeq \Rightarrow$ monotone $u(\cdot)$)

If $\succeq$ is monotone, then $\exists$ a monotone $u(\cdot)$ that represents $\succeq$.

Note *Complete, transitive, and monotone are three assumptions that made by all theories (either EU or non-EU).*

Definition 1.6 (Strongly Monotone $\succeq$)

$\succeq$ is **strongly monotone** if and only if for any $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in X$, if $\forall i : x_i \geq y_i$ and $\exists j$ such that $x_j > y_j$, then $x \succ y$.

(When we compare elements that have more than one dimension, strongly monotone holds if at least one relation is not equal.)

$$A = (1,1), B = (2,1), C(1,2), D = (2,2)$$

Strongly monotone can infer that $D \succ B \succ A, D \succ C \succ A$.

Even weaker assumptions will ensure that the consumer's choice exhausts their budget.

Definition 1.7 (Local Nonsatiation)

For any bundle x, there is a nearby bundle y in the consumption set such that y is preferred to x. That is, for all $x \in X$ and every $\varepsilon > 0$,

$$\exists y \in |x - y| < \varepsilon, \text{ s.t. } y \succ x$$

♣

We have

$$\text{Strong Monotonicity} \Rightarrow \text{Monotonicity} \Rightarrow \text{Local Nonsatiation}$$

1.1.4 Independence of Preference

The 'standard' model of decisions under risk is based on von Neumann and Morgenstern Expected Utility (EU), which requires the independence assumption.

Definition 1.8 (Independence of Preference)

Independence: For any $x, y, z \in X$ and $0 < \alpha < 1$, if $x \succeq y$ then $\alpha x + (1-\alpha)z \succeq \alpha y + (1-\alpha)z$. ♣

1.1.5 Continuous Preference

Definition 1.9 (Continuous $\succeq$)

$\succeq$ is **continuous** on X if and only if for any sequence $\{x^n, y^n\}_n = 1^\infty$ with $x^n \succeq y^n$ and we note $x = \lim_{n\to\infty} x^n$, $y = \lim_{n\to\infty} y^n$, we have $x \succeq y$ (i.e., the graph $\{(x,y) \mid x \succeq y \subseteq X \times X\}$ is closed). ♣

Example 1.1 Lexicographic preferences (not continuous) Under Lexicographic preference $\succ$, $x \succ y$ if and only if

- $x_1 > y_1$, or
- $x_1 = y_1$, and $x_2 > y_2$, or
- $x_1 = y_1$ and $x_2 = y_2$ and $x_3 > y_3$, or
- etc.

Under Lexicographic preferences, there is no indifference.

We can find the Lexicographic preference violates continuity: $\left(1 + \frac{1}{n}, 1\right) \succ (1,2)$ and $\lim \left(1 + \frac{1}{n}, 1\right) = (1,1) \prec (1,2)$.

Proposition 1.3 (Debreu's Theorem, Continuous $\succeq \Rightarrow$ continuous $u(\cdot)$)

If $\succeq$ is continuous, then $\exists$ a continuous $u(\cdot)$ that represents $\succeq$.

1.1.6 Convex Preference

Definition 1.10 (Convex $\succeq$)

$\succeq$ is **convex** if for every $x \in X$ the $\{y \in X : y \succeq x\}$ is convex, i.e., $y_1 \succeq x$ and $y_2 \succeq x \Rightarrow \alpha y_1 + (1-\alpha)y_2 \succeq x$ for all $\alpha \in [0,1]$.

Convex relations imply *averages are preferred to extremes*.

Definition 1.11 (Strictly Convex)

$\succeq$ is **strictly convex** iff $\forall x, y, z \in X$, if $x \succeq z$ and $y \succeq z$, then $\alpha x + (1-\alpha)y \succ z$ for all $\alpha \in (0,1)$

Definition 1.12 (Quasi-Concave Function)

A function u is **quasi-concave** if and only if for all $y \in X$, $\{x \in X : u(x) \geq u(y)\}$ is convex. Any function that is concave is also quasi-concave.

Proposition 1.4 (Concave Function $\Rightarrow$ Quasi-Concave Function)

Any function that is concave is also quasi-concave.

Proposition 1.5 (Convex $\succeq \Leftrightarrow$ quasi-concave $u(\cdot)$)

$\succeq$ is convex, $\Leftrightarrow \exists$ a quasi-concave $u(\cdot)$ that represents $\succeq$.

1.1.7 Homothetic Preference

Definition 1.13 (Homotheticity)

$\succeq$ are homothetic if $x \succeq y \Rightarrow \alpha x \succeq \alpha y$ for all $\alpha > 0$.

Proposition 1.6 (Homothetic preference $\Leftrightarrow$ homogeneous $u(\cdot)$)

A continuous $\succeq$ is homothetic $\Leftrightarrow \exists$ a continuous homogeneous $u(\cdot)$ that represents $\succeq$ such that $u(\alpha x) = \alpha u(x)$ for all $x > 0$.

1.1.8 Quasi-linearity

Definition 1.14 (Quasi-Linearity)

$\succeq$ on X is **quasi-linear** on x_1 if

$$x \succeq y \Rightarrow (x + \epsilon e_1) \succeq (y + \epsilon e_1)$$

where $e_1 = (1, 0, ..., 0)$ and $\epsilon > 0$.

Theorem 1.1 (Quasi-Linearity $\Leftrightarrow u(x) = x_1 + v(x_{-1})$)

A continuous $\succeq$ on $(-\infty, \infty) \times \mathbb{R}_+^{K-1}$ is quasi-linear in $x_1 \Leftrightarrow \exists$ a $u(\cdot)$ that represents $\succeq$ such that

$$u(x) = x_1 + v(x_{-1})$$

where $v(\cdot)$ satisfies $(v(x_{-1}), 0, ..., 0) \sim (0, x_{-1})$.

1.1.9 Separability

Definition 1.15 (Separability)

$\succeq$ satisfies **separability** if for any x_i

$$(x_i, x_{-i}) \succeq (x_i', x_{-i}) \Leftrightarrow (x_i, x_{-i}') \succeq (x_i', x_{-i}')$$

Theorem 1.2 (Separability $\Rightarrow$ Additive $u(\cdot)$)

*$\succeq$ with **separability** admits additive u-representation*

$$u(x) = v_1(x_1) + \cdots v_K(x_K)$$

Note *Strong assumption, usually ignored in practice.*

1.1.10 Differentiable Preference

Consider a vector of values $v(x) \in \mathbb{R}_+^K$ for the K commodities and a feasible the direction $x + \varepsilon d \in X$ from x for small enough $\varepsilon > 0$.

d is considered improvement if and only if

$$d \cdot v(x) > 0$$

Given $v(x) : X \to \mathbb{R}_+^K$, let

$$D_v(x) = \{d : d \cdot v(x) > 0\}$$

be the set of directions that are improvements relative to x.

1.2 Choice

1.2.1 Choice

Definition 1.16

A **choice function** C such that $C(A) \subseteq A$ which specifies for each nonempty subset $A \subseteq X$.

A choice function C can be rationalized if there is a preference relation $\succeq$ on X such that $C = C_{\succeq}$, that is

$$C(A) = C_{\succeq}(A) = \{x \in A : x \succeq y, \forall y \in A\}, \forall A \subseteq X$$

Definition 1.17 (Rubinstein's Condition α)

A choice function C satisfies **condition** α if for any two problems A, B, if $A \subseteq B$ and $C(B) \in A$, then $C(A) = C(B)$. ♣

1.2.2 Choice Correspondence

Definition 1.18 (Choice Correspondence (More than one choice))

A choice correspondence C assigns a non-empty subset for every non-empty set A

$$\emptyset \neq C(A) \subseteq A$$

♣

Properties:

(α): If $a \in A \subseteq B$, then $a \in C(B) \Rightarrow a \in C(A)$

(β): If $a, b \in A \subseteq B$, then $a, b \in C(A)$ and $b \in c(B) \Rightarrow a \in C(B)$.

α and β are equivalent to WARP.

1.2.3 Weak Axiom of Revealed Preference (WARP)

Definition 1.19 (Weak Axiom of Revealed Preference)

Given a choice structure $(C, \mathcal{B})$ satisfies **WARP**. If $\exists B \in \mathcal{B}$ with $x, y \in B$, such that $x \in C(B)$. Then, $\forall B' \in \mathcal{B}$ with $x, y \in B'$, $y \in C(B') \Rightarrow x \in C(B')$.

Or we can say, whenever $x, y \in B \cap B'$,

$$x \in C(B) \text{ and } y \in C(B)' \Rightarrow x \in C(B')$$

♣

Proposition 1.7 (Rational $\Rightarrow$ WARP)

Given $\succeq$ is rational, then $(C^*_{\succeq}, \mathcal{B})$ satisfies WARP.

($C^*_{\succeq}$ is the choice rule that picks the maximal alternatives by $\succeq$) ♠

1.3 Social Choice

Notations:

1. We consider finite set of alternatives X and finite set of agents I.
2. We use $\mathcal{B}$ to denotes the set of all preference relations.
3. We use $\mathcal{R} \subseteq \mathcal{B}$ to denotes the set of all rational preference relations.

4. We use $\succeq \in \mathcal{R}$ to represents individual rational preference relation.

1.3.1 Social Welfare Function and Properties

Definition 1.20 (Social Welfare Function (SWF))

A **social welfare function** (SWF) is a mapping

$$f : \mathcal{A} \subseteq \mathcal{R}^I \to \mathcal{B}$$

$\trianglerighteq = f(\succeq_1, ..., \succeq_I)$ is interpreted as the **social preference relation**. It doesn't need to be rational (i.e., complete and transitive). ♣

Definition 1.21 (SWF's Properties)

A social welfare function $f : \mathcal{A} \to \mathcal{B}$

- has **unrestricted domain** (UD) if $\mathcal{A} = \mathcal{R}^n$;
- is **transitive** (T) if $f(\succeq_1, ..., \succeq_I)$ is transitive for all $(\succeq_1, ..., \succeq_I) \in \mathcal{A}$;
- is **nondictatorial** (ND) if there is no agent $i \in I$ such that $\forall \{x, y\} \subseteq X$ $x \succeq_i y \Rightarrow x \trianglerighteq y$. (That is there is no distinguished voter who can choose the winner).
- is **weakly Paretian** (PA) if, $\forall \{x, y\} \subseteq X$ and any preference profile $(\succeq_1, ..., \succeq_I) \in \mathcal{A}$, we have $x \succeq_i y, \forall i \in I \Rightarrow x \trianglerighteq y$.
- is **independent of irrelevant alternatives** (IIA) if, $\forall \{x, y\} \subseteq X$, and any $\succeq$ and $\succeq'$ with $\succeq_i|_{x,y} = \succeq'_i|_{x,y}, \forall i \in I$, if $x \trianglerighteq y$ then $x \trianglerighteq' y$. ♣

1.3.2 Arrow's Theorem

Theorem 1.3 (Arrow's impossibility theorem)

Suppose $|X| \geq 3$, $\mathcal{A} = \mathcal{R}^I$ (UD). Then if a SWF f satisfies T, PA, and IIA, then it fails to be ND. ♡

Proof 1.1

Yu, N. N. (2012). A one-shot proof of Arrow's impossibility theorem. *Economic Theory*, 523-525.

1.4 Demand Theory and Equilibrium

Budget set is given by $B = \{x \in X \subseteq \mathbb{R}_+^K : p \cdot x \leq w\}$, where w is the DM's wealth and p is the vector of prices. Without losing generality, we can assume $w = 1$.

The DM's problem is finding the $\succeq$-best bundle $x \in B(p)$.

Lemma 1.1

If $\succeq$ is continuous, then all such problems have a solution. ♡

Lemma 1.2

If $\succeq$ is convex, then the set $\succeq$-best bundle $x \in B(p)$ is convex.

Lemma 1.3

If $\succeq$ is strictly convex, then the set $\succeq$-best bundle $x \in B(p)$ is (at most) a singleton.

Assume that $\succeq$ is differentiable and denote the vector of "subjective value numbers" at x^* (as defined above) by $v(x^*) = (v_1(x^*), ..., v_K(x^*))$. If $x^* \in B()$

Consider a consumer's problem

$$\max_{x \in X} u(x)$$
$$s.t. \ p \cdot x \leq w \tag{UMP}$$

The set of all optimal solutions are represented by $x(p, w)$.

Proposition 1.8

If $p >> 0$ and $u(\cdot)$ is continuous, then UMP has a solution.

Solution: Marshallian (Uncompensated) Demand.

Proposition 1.9

If $\succeq$ is monotone, then $p \cdot x = w$ for all $x \in x(p, w)$.

Proposition 1.10

If $\succeq$ is convex, then the set of solutions $x(p, w)$ is convex.

Proof 1.2

Suppose $x, x' \in X$. The optimal utility $u^* = u(x) = u(x')$. For any $\alpha \in [0, 1]$, let $x'' = \alpha x + (1 - \alpha)x'$. Because $\succeq$ is convex, we have $u(\cdot)$ is quasi-concave, that is $u(x'') \geq u^*$. x'' is also feasible. So, $x'' \in x(p, w)$.

Consider the duality

$$\min_{x \in X} p \cdot x$$
$$s.t. \ u(x) \geq u \tag{EMP}$$

The optimal solutions are represented by $h(p, u)$.

Solution: Hicksian (compensated) demand.

Proposition 1.11

$u(\cdot)$ is monotone and $p >> 0$.

(i). For $w > 0$, if $x^* \in x(p, w)$, then $x^* \in h(p, u(x^*))$ and $px^* = w$.

(ii). For $u > u(0)$, if $x^* \in h(p, u)$, then $x^* \in x(p, p \cdot x^*)$ and $u(x^*) = u$.

Slutsky: how change of price in good k affects the demand of product l.

$$\frac{\partial x_l(p,w)}{\partial p_k} = \underbrace{\frac{\partial h_l(p,u)}{\partial p_k}}_{\text{substitution effect}} \underbrace{-\frac{\partial x_l(p,w)}{\partial w} x_k(p,w)}_{\text{income effect}}$$

Definition 1.22

Given endowment $\{w^i\}_{i\in I}$. A **competitive equilibrium** is a pair $p \in \mathbb{R}^L$ (price vector over L goods) and an allocation $(x^i)_{i\in I}$ such that:

(i). $x^i \in \operatorname{argmax} u^i(x)$ s.t. $p \cdot x^i \leq p \cdot w^i, \forall i \in I$.

(ii). $\sum_{i\in I} x^i_\rho(p, w) = \sum_{i\in I} w^i_\rho, \forall \rho \in L$.

Definition 1.23

An allocation x is **Pareto-effecient** if there doesn't exist an allocation y s.t. $u_i(y) \geq u_i(x)$ for each $i \in I$ and $u_j(y) > u_j(x)$ for some $j \in I$.

(Assume consumers' preferences are strict monotone)

Theorem 1.4 (First-order fundamental welfare theorem)

Suppose (p^*, x^*) is a competitive equilibrium. Then x^* is Pareto-efficient.

If CE exists we can prove a Pareto-effecient allocation is CE.

Theorem 1.5 (Second-order fundamental welfare theorem)

Suppose that x^* is Pareto efficient and consumers receive endowment worth $w^i = p \cdot x^{i*}$ for all $i \in I$. Then, if a competitive equilibrium exists for such w, then x^* is a competitive equilibrium allocation.

1.5 Basic Game Theory

1.5.1 von Neumann and Morgenstern Expected Utility (EU)

Definition 1.24 (von Neumann and Morgenstern Expected Utility (EU))

There a lottery $\mathcal{L}$ whose outcomes are denoted by $1, ..., N$ with probabilities $p_1, ..., p_N$ All lotteries have Bernoulli utility $u_1, ..., u_N$ for an agent.

With VNM (expected) utility, we can say two lotteries $\mathcal{L} \succeq \mathcal{L}'$ if and only if $\sum_{i=1}^N u_i p_i \geq \sum_{i=1}^N u_i p_i'$ ♣

1.5.2 Game, Dominant Strategy

A game is denoted by

$$\Gamma = \left(\underbrace{I}_{\text{players}}, \underbrace{\{S_i\}_{i \in I}}_{\text{Strategy Set}}, \underbrace{\{u_i(\cdot)\}_{i \in I}}_{\text{VNM utility}} \right)$$

$u_i : \prod_{i \in I} S_i \to \mathbb{R}$ is the utility function that maps all players' strategies to a player's utilities.

Definition 1.25 (Dominant Strategy)

A strategy $s_i \in S_i$ is **dominant** for i in Γ if for all $s_i' \neq s_i$, we have $u_i(s_i, S_{-i}) \geq u_i(s_i', S_{-i})$. ♣

1.5.3 Nash Equilibrium and Existence

Definition 1.26 (Nash Equilibrium)

A strategy profile $\Sigma = (\sigma_1, ..., \sigma_I)$ is a **Nash** equilibrium of the game Γ if for every $i \in I$, we have: $u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i', \sigma_{-i}^*), \forall \sigma_i' \in \Delta(S_i)$ (can't benefit from deviating). ♣

Theorem 1.6 (Existence of Nash Equilibrium)

A Nash equilibrium exists in Γ if for all $i \in I$,

(i). S_i is non-empty, convex, compact, subset of $\mathbb{R}^m$ (i.e., for some finite dimensions of real numbers).

(ii). $u_i(s_i, ..., s_I)$ is continuous in $(s_i, ..., s_I)$ and quasi-concave in any s_i. ♡

Proof 1.3

We prove a lemma for the best response correspondence $b_i(s_{-i}) = \operatorname{argmax}_{s_i \in S_i} u(s_i, s_{-i})$ firstly.

Lemma 1.4

Suppose $\{S_i\}_{i \in I}$ are non-empty. Suppose that S_i is compact and convex and u_i is continuous in $(s_i, ..., s_I)$ and quasi-concave in any s_i, then $b_i(s_{-i})$ is non-empty, convex-valued and uhc. ♡

Proof 1.4

This lemma is proved by Berge's Maximum Theorem (Theroem 7.2).

Consider the function $b : S \to S$ *with* $b(s_i, ..., s_I) = \{b_1(s_{-1}), ..., b_I(s_{-I})\}$.

As we proved b *is non-empty, convex-valued and uhc from* S *to* S *where* S *is non-empty, compact, and convex. By the Kakutani's Fixed Point Theorem (Theorem 7.6), we have* b *has a fixed point* $s \in S$*, which should be the Nash equilibrium.*

1.5.4 Bayesian Game

Definition 1.27 (Bayesian Game)

A **Bayesian game** is defined by $\Gamma = (I, \{S_i\}_{i\in I}, \{u_i(\cdot)\}_{i\in I}, \{\Theta_i\}_{i\in I}, \{F_i\}_{i\in I})$ where $u_i(s_i, s_{-i}, \theta_i)$ maps the strategies of players and player i's type $\theta_i \in \Theta_i$ to player i's utilities. F_i is the distribution of the player i's type. ♣

1.5.5 Mechanism Design

Design incentives for agents to reveal their types or achieve particular society outcomes.

Given the set of agents, alternatives (for the society), and types (of agents) are I, X, Θ and a social choice function $f : \Theta = (\Theta_1, ..., \Theta_I) \to X$.

Definition 1.28 (Mechanism)

A **mechanism** is represented as

$$\Gamma = \Big((S_1, ..., S_I), g : S \triangleq (S_1, ..., S_I) \to X\Big)$$

where S_i represents the strategy set of agent i. ♣

Definition 1.29 (Implement)

Γ (indirectly) **implements** a social choice function f if $\exists\, (s_1^*(\cdot), ..., s_I^*(\cdot))$ of a game induced by Γ such that $g(s_1^*(\theta_1), ..., s_I^*(\theta_I)) = f(\theta_1, ..., \theta_I)$ for all $(\theta_1, ..., \theta_I) \in \Theta$ ♣

Definition 1.30 (Direct Mechanism)

A mechanism is **direct mechanism** if $S_i = \Theta_i$ for all $i \in I$ and $g(\theta) = f(\theta)$ for all $\theta = (\theta_1, ..., \theta_I) \in \Theta$. So, a direct mechanism can be represented by $\Gamma = (\Theta, f(\cdot))$. ♣

Definition 1.31 (Weak Dominance)

A strategy is weakly dominant if for all $\theta_i \in \Theta_i$ and all $s_{-i}(\cdot) \in S_{-i}$, we have:

$$u_i(g(s_i(\theta_i), s_{-i}), \theta_i) \geq u_i(g(s_i', s_{-i}), \theta_i)$$

for all $s_i' \neq s_i$. ♣

Definition 1.32 (Dominant Strategy Equilibrium)

Strategy profile $s^* = (s_1^*(\cdot), ..., s_I^*(\cdot))$ is a **dominant strategy (D-S) equilibrium** of $\Gamma = (S, g(\cdot))$ if for all $i \in I$ and $\theta_i \in \Theta$, we have:

$$u_i(g(s_i^*(\theta_i), s_{-i}), \theta_i) \geq u_i(g(s_i', s_{-i}), \theta_i)$$

for all $s_i' \in S_i$ and $s_{-i} \in S_{-i}$. ♣

Definition 1.33 (Implement in dominant strategies)

Γ **implements** f in **dominant strategies** if $\exists$ a dominant strategy (D-S) equilibrium s^* of Γ such that $g(s^*(\theta)) = f(\theta)$. ♣

Definition 1.34 (Strategy-Proof, DSIC)

f is **strategy-proof** (also called dominant-strategy-incentive-compatible, DSIC) if "$s_i^*(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$ and all $i \in I$" is a dominant strategy (D-S) equilibrium of the direct mechanism $\Gamma = (\Theta, f(\cdot))$. ♣

Theorem 1.7 (Revelation Principle)

Suppose that $\exists \Gamma = (S, g(\cdot))$ that (indirectly) implements f in dominant strategies. Then f is strategy-proof (DSIC). ♡

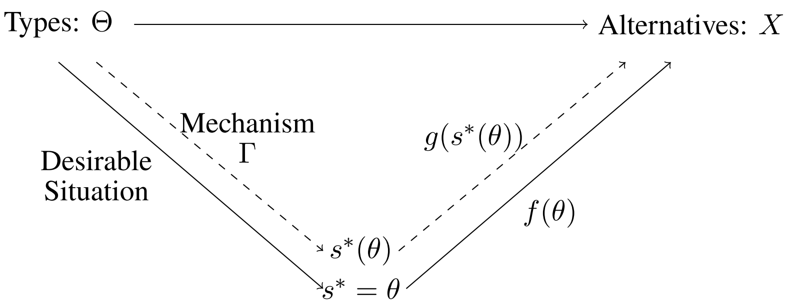

Figure 1.2: How Mechanism Design works

Theorem 1.8 (Gibbard-Satterthwaite theorem)

Suppose that $|X| \geq 3$ and a social choice function f is surjective. Then,

$$f \text{ is strategy-proof (DSIC)} \Leftrightarrow f \text{ is dictatorial (1.21)}$$ ♡

Some lemmas can help to prove the theorem.

Lemma 1.5

If f is strategy-proof (DSIC) and $f(\succeq) = x$ and $x \succeq_i y \Rightarrow x \succeq'_i y$ for all $i \in I$ and all $x \neq y \in X$, then $f(\succeq') = x$.

Lemma 1.6 (Pareto Effeciency)

If f is strategy-proof (DSIC) and $x \succ_i y$ for all $i \in I$, then $f(\succeq') \neq y$.

Example 1.2 Define $\succeq = \begin{pmatrix} x & y \\ y & x \\ z & z \end{pmatrix}$ and $\succeq' = \begin{pmatrix} x & y \\ y & z \\ z & x \end{pmatrix}$, each column 1/2 represents player 1/2's preferences.

Chapter 2 Market Design

Based on

- Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis, Roth, Alvin E.& Sotomayor, Matilda, 1990.
- Fleiner, T. (2003). A fixed-point approach to stable matchings and some applications. *Mathematics of Operations research*, 28(1), 103-126.
- Hatfield, J. W., & Kominers, S. D. (2017). Contract design and stability in many-to-many matching. *Games and Economic Behavior*, 101, 78-97.

2.1 Matching One-to-One

Suppose there are doctors (D) and hospitals (H). For a doctor d, define a relation $\succeq_d$ over $H \cup \{d\}$; for a hospital h, define a relation $\succeq_h$ over $D \cup \{h\}$. A matching market is defined by

$$(D, H, \{\succeq_i\}_{i \in D \cup H})$$

Note *Given a matching $\mu : D \cup H \to D \cup H$, we would call $\mu(d)$ be "d's match".*

Definition 2.1 (Involution)

A matching $\mu : D \cup H \to D \cup H$ is **involution** such that

$$\mu(d) \neq d \Rightarrow \mu(d) \in H, \forall d \in D$$

and

$$\mu(h) \neq h \Rightarrow \mu(h) \in D, \forall h \in H$$

♣

Definition 2.2 (Stable)

A matching $\mu : D \cup H \to D \cup H$ is **stable** if it is

- Individually Rational: $\nexists\, i$ for whom $i > \mu(i)$.
- (Pairwise) Unblocked: $\nexists\, (d, h)$ such that $d \succ_h \mu(h)$ and $h \succ_d \mu(d)$.

♣

Theorem 2.1 (Gale-Shapley, 1962)

For any matching market, a stable matching μ exists.

♡

Proof 2.1

We prove it by an algorithm:

Definition 2.3 (Deferred Acceptance Algorithm (DA))

At each round, every doctor applies for his most preferred hospital among those haven't rejected him. Each hospital chooses its most preferred doctors among its applicants and the one on the previous waitlist, and then rejects others.

Observation: DA terminates μ. We want to prove

1. μ *is IR (obviously);*
2. μ *is unblocked.*

   Suppose there is a block (d, h) *such that* $d \succ_h \mu(h)$ *and* $h \succ_d \mu(d)$. *That is impossible, because the* $d \neq \mu(h)$, *the* d *must be rejected by* h, *which means* $h \preceq_d \mu(d)$.

Note *We call "*h *is* ***achievable*** *for* d*" if* $\mu(d) = h$ *for some stable matching* μ.

2.1.1 Matching Markets: One-to-One

Definition 2.4 (D-Optimal Matching)

A matching $\mu : D \cup H \to D \cup H$ is D-**optimal**, denoted by μ^D, if for any stable μ' we have that $\mu^D \succeq_D \mu'$ (the best stable matching for all doctors).

Theorem 2.2 (Deferred Acceptance Algorithm $\Rightarrow$ D-Optimal Matching)

Deferred Acceptance Algorithm (with D proposing) terminates in μ^D.

Proof 2.2

...Theorem 2.12 (Gale and Shapley)

Theorem 2.3 (Lone-Wolf Theorem)

The set of matched agent is identical in every stable μ.

Proof 2.3

$|\mu^D(H)| \geq |\mu(H)| \geq |\mu^H(H)|$; *by symmetry,* $|\mu^H(D)| \geq |\mu(D)| \geq |\mu^D(D)|$. *Because* $|\mu^D(H)| = |\mu^D(D)|$ *and* $|\mu^H(H)| = |\mu^H(D)|$ *by one-to-one, so everything is equal.*

2.1.2 Joint and Meet

Definition 2.5 (Joint and Meet)

1. **Join** $\mu \vee_D \mu'$ assign the more preferred match to every d and the less preferred match to every h, that is,

$$\mu \vee_D \mu'(d) = \begin{cases} \mu(d), & \text{if } \mu(d) >_d \mu'(d) \\ \mu'(d), & \text{otherwise} \end{cases}, \forall d \in D$$

$$\mu \vee_D \mu'(h) = \begin{cases} \mu(h), & \text{if } \mu(h) <_h \mu'(h) \\ \mu'(h), & \text{otherwise} \end{cases}, \forall h \in H$$

2. **Meet** $\mu \wedge_D \mu'$ assign the less preferred match to every d and the more preferred match to every h, that is,

$$\mu \wedge_D \mu'(d) = \begin{cases} \mu(d), & \text{if } \mu(d) <_d \mu'(d) \\ \mu'(d), & \text{otherwise} \end{cases}, \forall d \in D$$

$$\mu \wedge_D \mu'(h) = \begin{cases} \mu(h), & \text{if } \mu(h) >_h \mu'(h) \\ \mu'(h), & \text{otherwise} \end{cases}, \forall h \in H$$

♣

Theorem 2.4 (Join and Meet of Stable Matchings are Stable)

If μ and μ' are stable, then $\mu \vee_D \mu'$ and $\mu \wedge_D \mu'$ are stable. ♡

2.1.3 Strategic Incentives

- Type $=$ preference list.
- SCF: $f : \Theta \to \mathcal{M}$, where $\mathcal{M}$ is a set of stable matchings;
- Is f strategy-proof?
- Does there exist a stable and strategy-proof (direct) mechanism?

Definition 2.6

We say a mechanism φ is strategy-proof (SP) if $\varphi(\succ_i, \succ_{-i}) \geq \varphi(\succ_i', \succ_{-i})$ for all $i \in I$ and $\succ_i'$ and $\succ_{-i}$. ♣

Theorem 2.5 (Impossibility theorem (Roth))

There is no stable and strategy-proof (SP) mechanism. ♡

The mechanism that yields the D-optimal stable matching (in terms of the stated preferences) makes it a dominant strategy for each doctor to state his true preferences. (Similarly, the mechanism that yields the H-optimal stable matching makes it a dominant strategy for every hospital to state its true preferences.)

Theorem 2.6 (Dubins and Freedman; Roth)

The doctor(D)-optimal stable mechanism is strategy-proof for doctors.

Proof 2.4

2.2 Matching Many-to-Many

Contracts are denoted by $x \in X, x_D \in D, x_H \in H$. $F \triangleq D \cup H$.

Consider a set of contracts $Y \subseteq X$,

- Y_D = doctors listed in Y;
- Y_d = the contract in Y that list the doctor d;
- $\succ_d$ over set of contracts that name the doctor d;
- The set of contracts $f \in F$ chooses from Y: $C_f(Y) = \max_{\succ_f}\{Z \subseteq X : Z \subseteq Y_f\} \subseteq Y_f$;
- The set of contracts doctors choose from Y: $C_D(Y) = \cup_{d \in D} C_d(Y)$.
- The set of contracts doctors reject from Y: $R_D(Y) = Y \backslash C_D(Y)$.

The outcome of matching is $Y \subseteq X$.

Definition 2.7 (Stable Contracts)

$A \subseteq X$ is **stable** if

- Individually Rational (IR): for all $f \in F$: $C_f(A) = A_f$;
- Unblocked: $\nexists$ non-empty $Z \subseteq X$ such that $Z \cap A = \emptyset$ and for all $f \in F$, $Z_f \subseteq C_f(A \cup Z)$.

Example 2.1 Preferences over doctor d: $\{x, y\} > \{x\} > \emptyset > \{y\}$; Preferences over hospital h: $\{y\} > \{x, y\} > \{x\} > \emptyset$.

$\{x\} \Rightarrow \{x, y\} \Rightarrow \{y\} \Rightarrow \emptyset \Rightarrow \{x\}$.

Definition 2.8 (Substitutability Condition)

Preference of f satisfies the **substitutability condition** if for all $Y \subseteq X$ and $x, z \in X \backslash Y$:

$$z \notin C_f(Y \cup \{z\}) \Rightarrow z \notin C_f(Y \cup \{z\} \cup \{x\})$$

($Y' \subseteq Y \subseteq X \Rightarrow R_f(Y') \subseteq R_f(X)$, where R is the rejection choice.)

If z is rejected given a set, then it should also be rejected given a larger set.

Theorem 2.7

If contracts are substitutes, then $Y \subseteq X$ is stable <u>if and only if</u> pairwise stable.

Proof 2.5

Prove $\Leftarrow$: (If not pairwise stable $\Rightarrow$ not stable)

Suppose that Z is a block. So, $Z \subseteq C_f(A \cup Z)$ for all f listed in Z.

We can pick a $z \in Z$ such that $z \in C_f(A \cup Z)$. By the substitutability condition, $z \in C_f(A \cup \{z\})$. So, it is stable.

Theorem 2.8

If contracts are substitutes, then a stable outcome exists. ♡

Definition 2.9 (Lattice)

On a **lattice**, $L = (X, <, \wedge, \vee)$ (or we just use $L = (X, <)$), $<$ is a partial order on X in such a way that any two elements x and y of X have a unique greatest lower bound (glb) $x \wedge y$ (meet) and a unique lowest upper bound (lub) $x \vee y$ (join). ♣

Definition 2.10 (Complete Lattice)

A lattice $L = (X, <)$ is **complete** if there are both a meet (i.e. a greatest lower bound) and a join (i.e. a least upper bound) for any subset $Y \subseteq X$.

These generalized meet and join operations on Y are denoted by $\wedge Y$ and $\vee Y$. ♣

Definition 2.11 (Monotone Function over Lattice)

A function from one lattice to another lattice $f : (X, <) \to (X', <')$ is **monotone** if $x \leq y \Rightarrow f(x) \leq' f(y)$ for any $x, y \in X$. ♣

Theorem 2.9 (Tarski 1955)

Let $L = (X, <)$ be a complete lattice and $f : L \to L$ be monotone ($\leq$) function on L. Then, the set $\{x \in L : f(x) = x\}$ of fixed points is a non-empty, complete lattice with order $\leq$. ♡

Proof 2.6

Fleiner, T. (2003). A fixed-point approach to stable matchings and some applications. Mathematics of Operations research, 28(1), 103-126.

Chapter 3 Stochastic Dominance

Based on

- MIT 14.123 S15 Stochastic Dominance Lecture Notes
- Princeton ECO317 Economics of Uncertainty Fall Term 2007 Notes for lectures 4. Stochastic Dominance
- Jensen, M. K. (2018). Distributional comparative statics. *The Review of Economic Studies*, 85(1), 581-610.

3.1 General Definitions

Definition 3.1 (Jensen (2018), Defnition 1)

Let F and G be two distributions on the same measurable space. Let u be a function for which the following expression is well-defined,

$$\int u(x)dF \geq \int u(x)dG \tag{3.1}$$

Then:

- F **first-order stochastically dominates** G if 3.1 holds for any increasing function u.
- F is a **mean-preserving spread** of G if 3.1 holds for any convex function u.
- F is a **mean-preserving contraction** of G if 3.1 holds for any concave function u.
- F **second-order stochastically dominates** G if 3.1 holds for any concave and increasing function u.
- F **dominates** G **in the convex-increasing order** if 3.1 holds for any convex and increasing function u.

Note *F is a **mean-preserving contraction** of $G \Leftrightarrow G$ is a **mean-preserving spread** of F.*

Definition 3.2 (MPS and MPC)

We define the following notations of sets.

- MPS(f) is the set of all **mean-preserving spread** of f;
- MPC(f) is the set of all **mean-preserving contraction** of f;

3.2 First-order Stochastic Dominance

3.2.1 Two Equivalent Definitions

Definition 3.3 (First-order Stochastic Dominance)

For any lotteries F and G, F **first-order stochastically dominates** G if and only if the decision maker weakly prefers F to G under every weakly increasing utility function u, i.e.,

$$\int u(x)dF \geq \int u(x)dG$$

♣

Definition 3.4 (First-order Stochastic Dominance)

For any lotteries F and G, F **first-order stochastically dominates** G if and only if

$$F(x) \leq G(x), \forall x$$

♣

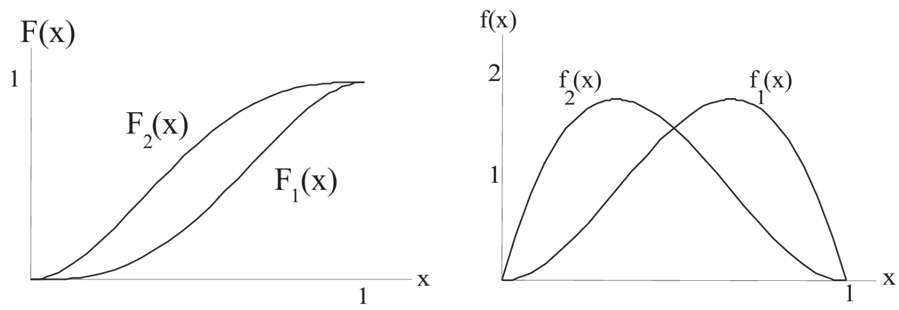

Figure 3.1: F_1 is FOSD over F_2: CDF and density comparison

3.3 Second-order Stochastic Dominance

3.3.1 Definition in terms of final goals

Definition 3.5 (Second-order Stochastic Dominance)

For any lotteries F and G, F **second-order stochastically dominates** G if and only if the decision maker weakly prefers F to G under every weakly **increasing concave** utility function u, i.e.,

$$\int u(x)dF \geq \int u(x)dG$$

♣

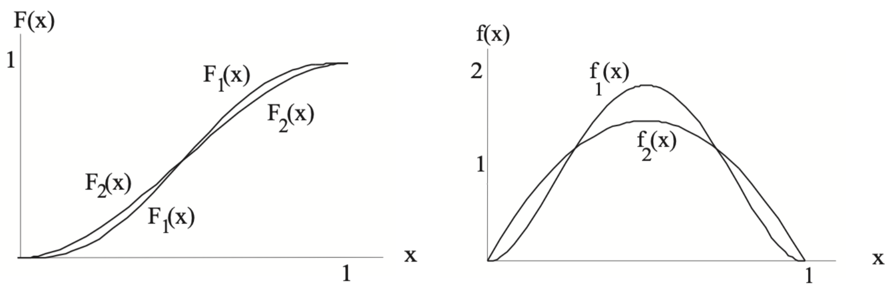

Figure 3.2: F_1 is SOSD over F_2: CDF and density comparison

3.3.2 Mean-Preserving Spread/Contraction

Definition 3.6 (Mean-Preserving Spread)

Let x_F and x_G be the random variables associated with lotteries F and G. Then G is a **mean-preserving spread** of F if and only if

$$x_G \stackrel{d}{=} x_F + \varepsilon$$

for some random variable ε such that $\mathbb{E}(\varepsilon \mid x_F) = 0, \forall x_F$.

The "$\stackrel{d}{=}$" means "is equal in distribution to" (that is, "has the same distribution as").

Note *Given G is a **mean-preserving spread** of F, G has larger variance than F.*

Example 3.1 $F(198) = \frac{1}{2}, F(202) = \frac{1}{2}$ and $G(100) = \frac{1}{100}, G(200) = \frac{98}{100}, G(300) = \frac{1}{100}$. Then

$$x_G \stackrel{d}{=} x_F + \varepsilon$$

where the distribution of ε can be solved by $\begin{cases} \frac{1}{100} = \frac{1}{2}P(\varepsilon = 102) + \frac{1}{2}P(\varepsilon = 98) \\ \frac{98}{100} = \frac{1}{2}P(\varepsilon = 2) + \frac{1}{2}P(\varepsilon = -2) \\ \frac{1}{100} = \frac{1}{2}P(\varepsilon = -98) + \frac{1}{2}P(\varepsilon = -102) \end{cases}$

3.3.3 For Same Mean Distributions, Second-order Stochastic Dominance is equivalent to Mean-Preserving Spread

Theorem 3.1 (Second-order Stochastic Dominance Equivalence)

Given $\int x dF = \int x dG$ (same mean). The following are equivalent.

1. *F second-order stochastically dominates G: $\int u(x) dF \geq \int u(x) dG$ for every weakly increasing concave utility function u.*
2. *F is a mean-preserving contraction of G (G is a mean-preserving spread of F).*
3. *For every $t \geq 0$, $\int_a^t G(x)dx \geq \int_a^t F(x)dx$.*

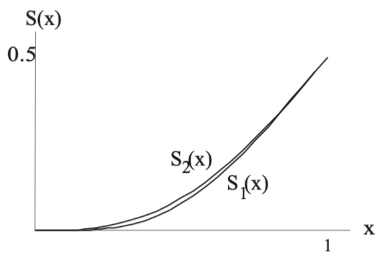

Figure 3.3: F_1 is SOSD over F_2, $S(t) : \int_a^t F_2(x)dx \geq \int_a^t F_1(x)dx$

Corollary 3.1 (Euqivalent Definitions of MPC and MPS)

F is a mean-preserving contraction of G (or G is a mean-preserving spread of F) if and only if

(1). $\int xdF = \int xdG$

(2). $\int_a^t G(x)dx \geq \int_a^t F(x)dx, \forall t$

♡

Corollary 3.2 (MPC(f) and MPS(f) are convex and compact)

MPC(f) *and* MPS(f) *are* ***convex*** *and* ***compact***.

♡

Chapter 4 Signalling Game

Based on

- "Kreps, D. M., & Sobel, J. (1994). Signalling. *Handbook of game theory with economic applications*, 2, 849-867."
-

4.1 Canonical Game

Definition 4.1 (Canonical Game)

1. There are two players: $\mathbf{S}$ (sender) and $\mathbf{R}$ (receiver).
2. $\mathbf{S}$ holds more information than $\mathbf{R}$: the value of some random variable t with support $\mathcal{T}$. (We say that t is the **type** of $\mathbf{S}$)
3. Prior belief of $\mathbf{R}$ concerning t are given by a probability distribution ρ over $\mathcal{T}$ (common knowledge)
4. $\mathbf{S}$ sends a **signal** $s \in \mathcal{S}$ to $\mathbf{R}$ drawn from a signal set $\mathcal{S}$.
5. $\mathbf{R}$ receives this signal, and then takes an **action** $a \in \mathcal{A}$ drawn from a set $\mathcal{A}$ (which could depend on the signal s that is sent).
6. $\mathbf{S}$'s payoff is given by a function $u : \mathcal{T} \times \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ and $\mathbf{R}$'s payoff is given by a function $v : \mathcal{T} \times \mathcal{S} \times \mathcal{A} \to \mathbb{R}$.

♣

4.2 Nash Equilibrium

Definition 4.2 (Strategy)

A **behavior strategy** for $\mathbf{S}$ is given by a function $\sigma : \mathcal{T} \times \mathcal{S} \to [0, 1]$ such that $\sum_s \sigma(t, s)$ for each t.

A **behavior strategy** for $\mathbf{R}$ is given by a function $\alpha : \mathcal{S} \times \mathcal{A} \to [0, 1]$ such that $\sum_a \alpha(s, a)$ for each t. ♣

Definition 4.3 (Nash Equilibrium)

Behavior strategies α and σ form a **Nash equilibrium** if and only if

1. For all $t \in \mathcal{T}$,

$$\sigma(t, s) > 0 \text{ implies } \textstyle\sum_a \alpha(s, a)u(t, s, a) = \max_{s' \in \mathcal{S}} \left(\sum_a \alpha(s', a)u(t, s', a)\right)$$

2. For each $s \in \mathcal{S}$ such that $\sum_t \sigma(t, s)\rho(t) > 0$,

$$\alpha(s, a) > 0 \text{ implies } \textstyle\sum_t \mu(t; s)v(t, s, a) = \max_{a'} \sum_t \mu(t; s)v(t, s, a')$$

where $\mu(t;s)$ is the $\mathbb{R}$'s posterior belief about t given s, $\mu(t;s) = \frac{\sigma(t,s)\rho(t)}{\sum_{t'}\sigma(t',s)\rho(t')}$ if $\sum_t \sigma(t,s)\rho(t) > 0$ and $\mu(t;s) = 0$ otherwise. ♣

Definition 4.4 (Separating & Pooling Equilibrium)

An equilibrium (σ, α) is called a **separating** equilibrium if each type t sends different signals; i.e., the set $\mathcal{S}$ can be partitioned into (disjoint) sets $\{\mathcal{S}_t; t \in \mathcal{S}\}$ such that $\sigma(t, \mathcal{S}_t) = 1$. An equilibrium (σ, α) is called a **pooling** equilibrium if there is a single signal s^* that is sent by all types; i.e., $\sigma(t, s^*) = 1$ for all $t \in \mathcal{T}$. ♣

4.3 Single-crossing

4.3.1 Situation over real line

Consider the situation that $\mathcal{T}, \mathcal{S}, \mathcal{A} \subseteq \mathbb{R}$ and $\geq$ is the usual "greater than or equal to" relationship.

1. We let $\Delta\mathcal{A}$ denote the set of probability distributions on $\mathcal{A}$.
2. For each $s \in \mathcal{S}$ and $\mathcal{T}' \subseteq \mathcal{T}$, we let $\Delta\mathcal{A}(s, T')$ be the set of mixed strategies that are the best responses by $\mathbf{R}$ to $s \in \mathcal{S}$ for some probability distribution with support $\mathcal{T}'$.
3. For $\alpha \in \Delta\mathcal{A}$, we write $u(t, s, \alpha) \triangleq \sum_{a\in\mathcal{A}} u(t, s, a)\alpha(a)$.

Definition 4.5 (Single-crossing)

The data of the game are said to satisfy the **single-crossing property** if the following holds: If $t \in \mathcal{T}$, $(s, \alpha) \in \mathcal{S} \times \Delta\mathcal{A}$ and $(s', \alpha') \in \mathcal{S} \times \Delta\mathcal{A}$ are such that $\alpha \in \Delta\mathcal{A}(s, \mathcal{T})$, $\alpha' \in \Delta\mathcal{A}(s', \mathcal{T})$, $s > s'$ and $u(t, s, \alpha) \geq u(t, s', \alpha')$, then for all $t' \in T$ such that $t' > t$, $u(t', s, \alpha) \geq u(t', s', \alpha')$. ♣

Chapter 5 Tools for Comparative Statics

Consider the function $f : (0, 2\pi) \times \mathbb{R} \to \mathbb{R}$ s.t.

$$f(x, a) = \sin x + a$$

Let $X = (0, 2\pi)$ and let $f_a(x) = f(x, a) = \sin x + a$ denote the perturbed function for fixed a.

5.1 Regular and Critical Points and Values

5.1.1 Rank of Derivatives $\text{Rank}df_x = \text{Rank}Df(x)$

Suppose $X \subseteq \mathbb{R}^n$ is open. Suppose $f : X \to \mathbb{R}^m$ is differentiable at $x \in X$, and let $W = \{e_1, ..., e_n\}$ denote the standard basis of $\mathbb{R}^n$. Then $df_x \in L(\mathbb{R}^n, \mathbb{R}^m)$, and

$$\begin{aligned}
\text{Rank}df_x &= \dim \text{Im}(df_x) \\
&= \dim \text{span}\{df_x(e_1), ..., df_x(e_n)\} \\
&= \dim \text{span}\{Df(x)e_1, ..., Df(x)e_n\} \\
&= \dim \text{span}\{\text{column 1 of } Df(x), ..., \text{column n of } Df(x)\} \\
&= \text{Rank}Df(x)
\end{aligned}$$

Thus,

$$\text{Rank}df_x \le \min\{m, n\}$$

df_x has **full rank** if $\text{Rank}df_x = \min\{m, n\}$, that is, is df_x has the maximum possible rank.

5.1.2 Regular and Critical Points and Values

Definition 5.1 (Regular and Critical Points and Values)

Suppose $X \subseteq \mathbb{R}^n$ is open. Suppose $f : X \to \mathbb{R}^m$ is differentiable at $x \in X$.

1. x is a **regular point** of f if $\text{Rank}df_x = \min\{m, n\}$.
2. x is a **critical point** of f if $\text{Rank}df_x < \min\{m, n\}$.
3. y is a **critical value** of f if there exists $x \in f^{-1}(y)$ such that x is a critical point of f.
4. y is a **regular value** of f if y is not a critical value of f.

Note *Notice that if $y \notin f(X)$, so $f^{-1}(y) = \emptyset$, then y is automatically a regular value of f.*

Example 5.1 Suppose $f(x,y) = (\sin x, \cos y)$, $Df(x,y) = \begin{bmatrix} \cos x & 0 \\ 0 & -\sin y \end{bmatrix}$. Critical point: $\{(\frac{k\pi}{2}, \mathbb{R}) : k \in 2\mathbb{Z}+1\} \cup \{(\mathbb{R}, k\pi) : k \in \mathbb{Z}\}$; Critical values: $\{(x,y) : x = 1$ or $x = -1$ or $y = 1$ or $y = -1\}$

5.2 Inverse and Implicit Function Theorem

5.2.1 Inverse Function Theorem

Using Taylor's theorem to approximate

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + o(x - x_0)$$

The requirement of "regular point" is necessary for the $Df(x_0)$ being invertible.

Theorem 5.1 (Inverse Function Theorem)

Suppose $X \subseteq \mathbb{R}^n$ is open. Suppose $f : X \to \mathbb{R}^n$ is C^1 on X, and $x_0 \in X$. If $\det Df(x_0) \neq 0$ (i.e., x_0 is a regular point of f), then there are open neighborhoods U of x_0 and V of $f(x_0)$ s.t.

$$f : U \to V \text{ is bijective (on-to-on and onto)}$$

$$\exists\, f^{-1} : V \to U \text{ is } C^1$$

$$Df^{-1}(f(x_0)) = [Df(x_0)]^{-1}$$

$$(\text{In } \mathbb{R},\ (f^{-1})'(f(x_0)) = (f'(x_0))^{-1})$$

If in addition $f \in C^k$, then $f^{-1} \in C^k$. ♡

5.2.2 Implicit Function Theorem

Using Taylor's theorem to approximate

$$f(x,a) = f(x_0,a_0) + Df(x_0,a_0)(x - x_0) + Df(x_0,a_0)(a - a_0) + \text{remainder}$$

The requirement of "regular point" is necessary for the $Df(x_0, a_0)$ being invertible.

We want to know how the function $x^*(a)$ changes with keeping $f(x^*, a) = 0$.

Theorem 5.2 (Implicit Function Theorem)

Suppose $X \subseteq \mathbb{R}^n$ and $A \subseteq \mathbb{R}^p$ are open and $f : X \times A \to \mathbb{R}^n$ is C^1. Suppose $f(x_0, a_0) = 0$ and $\det(D_x f(x_0, a_0)) \neq 0$, i.e. x_0 is a regular point of $f(\cdot, a_0)$. Then there are open neighborhoods U of x_0 ($U \subseteq X$) and W of a_0 such that

$$\forall a \in W,\ \exists! x \in U \text{ s.t. } f(x,a) = 0$$

For each $a \in W$ let $g(a)$ be that unique x. Then $g : W \to U$ is C^1 and

$$Dg(a_0) = -[D_x f(x_0, a_0)]^{-1}[D_a f(x_0, a_0)]$$

If in addition $f \in C^k$, then $g \in C^k$.

5.2.3 Prove Implicit Function Theorem Given Inverse Function Theorem

Proof 5.1

1. *Firstly, we prove "g is differentiable": The "change of a" incurs the value change:*

$$\begin{aligned} f(x_0, a_0 + h) &= f(x_0, a_0) + D_a f(x_0, a_0)h + o(h) \\ &= D_a f(x_0, a_0)h + o(h) \end{aligned}$$

Find a Δx such that the new x can let the value go back to 0, i.e., $f(x_0 + \Delta x, a_0 + h) = 0$. That is,

$$g(a_0 + h) = x_0 + \Delta x$$

To prove "g is differentiable", we want to prove "$\exists T \in L(A, X)$ s.t. $\Delta x = T(h) + o(h)$"

$$\begin{aligned} 0 &= f(x_0 + \Delta x, a_0 + h) \\ &= f(x_0, a_0) + D_x f(x_0, a_0 + h)\Delta x + D_a f(x_0, a_0)h + o(\Delta x) + o(h) \\ &= D_x f(x_0, a_0 + h)\Delta x + D_a f(x_0, a_0)h + o(\Delta x) + o(h) \end{aligned}$$

$$D_x f(x_0, a_0 + h)\Delta x = -D_a f(x_0, a_0)h + o(\Delta x) + o(h)$$

Because f is C^1 and the determinant is a continuous function of the entries of the matrix, $\det D_x f(x_0, a_0 + h) \neq 0$ for h sufficiently small, so

$$\Delta x = -[D_x f(x_0, a_0 + h)]^{-1} D_a f(x_0, a_0)h + o(\Delta x) + o(h)$$

$$\text{Since } f \in C^1, \Delta x = -[D_x f(x_0, a_0) + o(1)]^{-1} D_a f(x_0, a_0)h + o(\Delta x) + o(h)$$

$$\text{Since } f \in C^1, \Delta x = -[D_x f(x_0, a_0)]^{-1} D_a f(x_0, a_0)h + o(\Delta x) + o(h)$$

Hence, "g is differentiable" is proved and the derivative of g is $Dg(a_0) = -[D_x f(x_0, a_0)]^{-1}[D_a f(x_0, a_0)]$.

2. *Secondly, given the "g is differentiable", we can also compute the derivative by*

$$Df(g(a), a)(a_0) = 0$$

$$D_x f(x_0, a_0)Dg(a_0) + D_a f(x_0, a_0) = 0$$

$$Dg(a_0) = -[D_x f(x_0, a_0)]^{-1} D_a f(x_0, a_0)$$

Example 5.2 $f : \mathbb{R}^3 \to \mathbb{R}^2$, $f((3,-1,2)) = (0,0)$, $Df(3,-1,2) = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix}$. Then, let $(x_0, a_0) = (3,-1,2)$, where $x_0 = 3$ and $a_0 = (-1,2)$. Or, we can let $(x_0, a_0) = (3,-1,2)$, where $x_0 = (3,-1)$ and $a_0 = 2$.

5.2.4 Prove Inverse Function Theorem Given Implicit Function Theorem

Proof 5.2 (Prove Inverse Function Theorem Given Implicit Function Theorem)

Define $F : X \times \mathbb{R}^n$ s.t. $F(x,y) = y - f(x)$. Let $y_0 = f(x_0)$.

$$D_x F(x,y) = -Df(x),\ D_y F(x,y) = I_{n\times n}$$

According to the implicit function theorem, there are open sets $U \subseteq X$ and $V \subseteq \mathbb{R}^n$ such that $x_0 \in U$, $y_0 \in V$ and a function $g : V \to U$ differentiable at y_0 such that $F(g(y), y) = 0$ for all $y \in V$. So, $0 = F(g(y), y) = y - f(g(y))$, we have $f(g(y)) = y$, that is $g = f^{-1}$. $f : U \to V$ is bijective because it has inverse $g : V \to U$.

By the implicit function theorem, $g(y)$ is differentiable and

$$Df^{-1}(y_0) = Dg(y_0) = -[D_x F(x_0, y_0)]^{-1}[D_y F(x_0, y_0)] = [Df(x_0)]^{-1}$$

where $y_0 = f(x_0)$.

By the implicit function theorem, the $g = f^{-1}$ is C^k if f is C^k.

All in all, the inverse function theorem is proved.

5.2.5 Example: Using Implicit Function Theorem in Comparative Statics

Example 5.3 Let us consider a firm that produces a good y; it uses two inputs x_1 and x_2. The firm sells the output and acquires the inputs in competitive markets: The market price of y is p, and the cost of each unit of x_1 and x_2 are w_1 and w_2 respectively. Its technology is given by $f : \mathbb{R}^2_+ \to \mathbb{R}_+$, where $f(x_1, x_2) = x_1^a x_2^b$, $a + b < 1$. Its profits take the form

$$\pi(x_1, x_2; p, w_1, w_2) = p x_1^a x_2^b - w_1 x_1 - w_2 x_2$$

The firm selects x_1 and x_2 in order to maximize profits. **We aim to know how its choice of x_1 and x_2 is affected by a change in w_1.**

Assuming an interior solution, the first-order conditions of this optimization problem are

$$\frac{\partial \pi}{\partial x_1}(x_1^*, x_2^*; p, w_1, w_2) = pa(x_1^*)^{a-1}(x_2^*)^b - w_1 = 0$$
$$\frac{\partial \pi}{\partial x_2}(x_1^*, x_2^*; p, w_1, w_2) = pb(x_1^*)^{a}(x_2^*)^{b-1} - w_2 = 0$$

for some $(x_1, x_2) = (x_1^*, x_2^*)$.

Let us define

$$F(x_1^*, x_2^*; p, w_1, w_2) = \begin{bmatrix} pa(x_1^*)^{a-1}(x_2^*)^b - w_1 \\ pb(x_1^*)^a(x_2^*)^{b-1} - w_1 \end{bmatrix}$$

Jacobian matrices are

$$D_{(x_1,x_2)}F(x_1^*, x_2^*; p, w_1, w_2) = \begin{bmatrix} pa(a-1)(x_1^*)^{a-2}(x_2^*)^b & pab(x_1^*)^{a-1}(x_2^*)^{b-1} \\ pab(x_1^*)^{a-1}(x_2^*)^{b-1} & pb(b-1)(x_1^*)^a(x_2^*)^{b-2} \end{bmatrix}$$

$$D_{w_1}F(x_1^*, x_2^*; p, w_1, w_2) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

By the implicit function theorem, we can get

$$\begin{aligned} \begin{bmatrix} \frac{\partial x_1^*}{\partial w_1} \\ \frac{\partial x_2^*}{\partial w_1} \end{bmatrix} &= -[D_{(x_1,x_2)}F(x_1^*, x_2^*; p, w_1, w_2)]^{-1}[D_{w_1}F(x_1^*, x_2^*; p, w_1, w_2)] \\ &= [D_{(x_1,x_2)}F(x_1^*, x_2^*; p, w_1, w_2)]^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned}$$

5.2.6 Corollary: $a \to \{x \in X : f(x,a) = 0\}$ is lhc

Corollary 5.1

Suppose $X \subseteq \mathbb{R}^n$ and $A \subseteq \mathbb{R}^p$ are open and $f : X \times A \to \mathbb{R}^n$ is C^1. If 0 is a regular value of $f(\cdot, a_0)$, then the correspondence

$$a \to \{x \in X : f(x,a) = 0\}$$

*is **lower hemicontinuous** at a_0.*

5.3 Transversality and Genericity

5.3.1 Lebesgue Measure Zero

Definition 5.2 (Lebesgue Measure Zero)

Suppose $A \subseteq \mathbb{R}^n$. A has **Lebesgue measure zero** if for every $\varepsilon > 0$ there is a countable collection of rectangles $I_1, I_2, \dots$ such that

$$\sum_{k=1}^{\infty} \text{Vol}(I_k) < \varepsilon \text{ and } A \subseteq \cup_{k=1}^{\infty} I_k$$

Here by a rectangle we mean $I_k = \times_{j=1}^{n}(a_j^k, b_j^k) = \{x \in \mathbb{R}^n : x_j \in (a_j^k, b_j^k), \forall j\}$ for some $a_j^k < b_j^k \in \mathbb{R}$, and

$$\mathrm{Vol}(I_k) = \prod_{j=1}^{n} |b_j^k - a_j^k|$$

♣

Example 5.4

1. "Lower-dimensional" sets have Lebesgue measure zero. For example, $A = \{x \in \mathbb{R}^2 : x_2 = 0\}$
2. Any **finite** set has Lebesgue measure zero in $\mathbb{R}^n$.
3. **Finite Union** of sets that have Lebesgue measure zero has Lebesgue measure zero: If A_n has Lebesgue measure zero $\forall n$ then $\cup_{n\in N} A_n$ has Lebesgue measure zero.
4. Every **countable** set (e.g. $\mathbb{Q}$) has Lebesgue measure zero.
5. No open set in $\mathbb{R}^n$ has Lebesgue measure zero.

5.3.2 Sard's Theorem

Theorem 5.3 (Sard's Theorem)

Let $X \subseteq \mathbb{R}^n$ be open, and $f : X \to \mathbb{R}^m$ be C^r with $r \geq 1 + max\{0, n - m\}$. Then the set of all critical values of f has Lebesgue measure zero.

♡

5.3.3 Transversality Theorem

Theorem 5.4 (Transversality Theorem)

Let $X \subseteq \mathbb{R}^n$ and $A \subseteq \mathbb{R}^p$ be open, and $f : X \times A \to \mathbb{R}^m$ be C^r with $r \geq 1 + max\{0, n - m\}$. Suppose that 0 is a regular value of f (that is all (x, a) such that $f(x, a) = 0$ are regular points). Then,

1. *$\exists A_0 \subseteq A$ such that $A \backslash A_0$ has Lebesgue measure zero.*
2. *$\forall a \in A_0$, 0 is a regular value of $f_a = f(\cdot, a)$.*

♡

Example 5.5 $f : \mathbb{R}^4 \to \mathbb{R}^3$ s.t. $f(x, y, z, w) = (g(x) + y, z^3 + 1, w + x + y^2)$

Chapter 6 Fixed Point Theorem

6.1 Contraction Mapping Theorem (@ Lec 05 of ECON 204)

6.1.1 Contraction: Lipschitz continuous with constant < 1

Definition 6.1

Let (X, d) be a nonempty complete metric space. An operator is a function $T : X \to X$. An operator T is a **contraction of modulus** β if $\beta < 1$ and

$$d(T(x), T(y)) \leq \beta d(x, y), \forall x, y \in X$$

♣

A contraction shrinks distances by a *uniform* factor $\beta < 1$.

6.1.2 Theorem: Contraction $\Rightarrow$ Uniformly Continuous

Theorem 6.1 (Contraction $\Rightarrow$ Uniformly Continuous)

Every contraction is uniformly continuous.

♡

Proof 6.1

Let $\delta = \frac{\varepsilon}{\beta}$.

6.1.3 Blackwell's Sufficient Conditions for Contraction

Let X be a set, and let $B(X)$ be the set of all bounded functions from X to $\mathbb{R}$. Then $(B(X), \|\cdot\|_\infty)$ is a normed vector space.

(Notice that below we use shorthand notation that identifies a constant function with its constant value in $\mathbb{R}$, that is, we write interchangeably $a \in \mathbb{R}$ and $a : X \to \mathbb{R}$ to denote the function such that $a(x) = a, \forall x \in X$.)

Theorem 6.2 (Blackwell's Sufficient Conditions)

Consider $B(X)$ *with the sup norm* $\|\cdot\|_\infty$. *Let* $T : B(X) \to B(X)$ *be an operator satisfying*

1. *(monotonicity)* $f(x) \leq g(x), \forall x \in X \Rightarrow (Tf)(x) \leq (Tg)(x), \forall x \in X$
2. *(discounting)* $\exists \beta \in (0, 1)$ *such that for every* $a \geq 0$ *and* $x \in X$,

$$(T(f + a))(x) \leq (Tf)(x) + \beta a$$

Then T *is a contraction with modulus* β.

♡

Proof 6.2

Fix $f, g \in B(X)$. By the definition of the sup norm,

$$f(x) \leq g(x) + \|f - g\|_\infty \forall x \in X$$

Then

$$(Tf)(x) \leq (T(g + \|f - g\|_\infty))(x) \leq (Tg)(x) + \beta\|f - g\|_\infty \quad \forall x \in X$$

where the first inequality above follows from monotonicity, and the second from discounting. Thus

$$(Tf)(x) - (Tg)(x) \leq \beta\|f - g\|_\infty \quad \forall x \in X$$

Reversing the roles of f and g above gives

$$(Tg)(x) - (Tf)(x) \leq \beta\|f - g\|_\infty \quad \forall x \in X$$

Thus

$$\|T(f) - T(g)\|_\infty \leq \beta\|f - g\|_\infty$$

Thus T is a contraction with modulus β

6.2 Fixed Point Theorem (@ Lec 05 of ECON 204)

6.2.1 Fixed Point

Definition 6.2 (Fixed Point)

A **fixed point** of an operator T is element $x^* \in X$ such that $T(x^*) = x^*$. ♣

Definition 6.3 (Fixed Point of Function)

Let X be a nonempty set and $f : X \to X$. A point $x^* \in X$ is a **fixed point** of f if $f(x^*) = x^*$. ♣

Example 6.1 Let $X = \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$

1. $f(x) = 2x$ has fixed point: $x = 0$.
2. $f(x) = x$ has fixed points: $x \in \mathbb{R}$.
3. $f(x) = x + 1$ doesn't have fixed points.

6.2.2 ★ Contraction Mapping Theorem: contraction $\Rightarrow$ exist unique fixed point

Theorem 6.3 (Contraction Mapping Theorem)

Let (X, d) be a nonempty complete metric space and $T : X \to X$ a contraction with modulus $\beta < 1$. Then

1. *T has a unique fixed point x^*.*
2. *For every $x_0 \in X$, the sequence defined by*

$$\begin{aligned} x_1 &= T(x_0) \\ x_2 &= T(x_1) = T(T(x_0)) = T^2(x_0) \\ &\vdots \\ x_{n+1} &= T(x_n) = T^{n+1}(x_0) \end{aligned}$$

converges to x^*.

Note that the theorem asserts both the **existence** and **uniqueness** of the fixed point, as well as giving an **algorithm** to find the fixed point of a contraction.

Proof 6.3

Define the sequence $\{x_n\}$ as above. Then,

$$\begin{aligned} d(x_{n+1}, x_n) &= d(T(x_n), T(x_{n-1})) \\ &\leq \beta d(x_n, x_{n-1}) \\ &\leq \beta^n d(x_1, x_0) \end{aligned}$$

Then for any $n > m$,

$$\begin{aligned} d(x_n, x_m) &\leq d(x_1, x_0) \sum_{i=m}^{n-1} \beta^i \\ &< d(x_1, x_0) \sum_{i=m}^{\infty} \beta^i \\ &= \frac{\beta^m}{1-\beta} d(x_1, x_0) \to 0 \text{ as } m \to \infty \end{aligned}$$

Fixed $\varepsilon > 0$, we can choose $N(\varepsilon)$ such that $\forall n, m > N(\varepsilon)$,

$$d(x_n, x_m) < \frac{\beta^m}{1-\beta} d(x_1, x_0) < \varepsilon$$

Therefore, $\{x_n\}$ is Cauchy. Since (X, d) is complete, $x_n \to x^*$ for some $x^* \in X$.

Next we show that x^* is a fixed point of T.

$$\begin{aligned} T(x^*) &= T\left(\lim_{n\to\infty} x_n\right) \\ &= \lim_{n\to\infty} T(x_n) \text{ since } T \text{ is continuous} \\ &= \lim_{n\to\infty} x_{n+1} \\ &= x^* \end{aligned}$$

so x^* is a fixed point of T.

Finally, we show that there is at most one fixed point. Suppose x^* and y^* are both fixed points of T, so $T(x^*) = x^*$ and $T(y^*) = y^*$. Then

$$\begin{aligned} d(x^*, y^*) &= d(T(x^*), T(y^*)) \\ &\le \beta d(x^*, y^*) \\ \Rightarrow (1-\beta) d(x^*, y^*) &\le 0 \\ \Rightarrow d(x^*, y^*) &\le 0 \end{aligned}$$

So $d(x^*, y^*) = 0$, which implies $x^* = y^*$.

6.2.3 Conditions for Fixed Point's Continuous Dependence on Parameters

Theorem 6.4 (Continuous Dependence on Parameters)

Let (X, d) and (Ω, ρ) be two metric spaces and $T : X \times \Omega \to X$. For each parameter $\omega \in \Omega$ let $T_\omega : X \to X$ be defined by $T_\omega(x) = T(x, \omega)$.

Suppose (1). (X, d) is complete, (2). T is continuous in ω (that is $T(x, \cdot) : \Omega \to X$ is continuous for each $x \in X$), and (3). $\exists \beta < 1$ such that T_ω is a contraction of modulus β $\forall \omega \in \Omega$.

Then the fixed point function (about parameter ω) $x^* : \Omega \to X$ defined by $x^*(\omega) = T_\omega(x^*(\omega))$ is continuous.

6.3 Brouwer's Fixed Point Theorem (@ Lec 13 of ECON 204)

6.3.1 Simple One: One-dimension

Theorem 6.5

Let $X = [a, b]$ for $a, b \in \mathbb{R}$ with $a < b$ and let $f : X \to X$ be continuous. Then f has a fixed point.

Proof 6.4

Easily proved by Intermediate Value Theorem.

6.3.2 ★ Brouwer's Fixed Point Theorem: continuous function has fixed point over compact, convex set

Theorem 6.6 (Brouwer's Fixed Point Theorem)

Let $X \subseteq \mathbb{R}^n$ be nonempty, ***compact****, and* ***convex****, and let $f : X \to X$ be continuous. Then f has a fixed point.* ♡

Proof 6.5

Consider the case when the set X is the unit ball in $\mathbb{R}^n$.

Using a fact that "Let B be the unit ball in $\mathbb{R}^n$. Then there is no continuous function $h : B \to \partial B$ such that $h(x_0) = x_0$ for every $x_0 \in \partial B$", which is intuitive but hard to prove. (See *J. Franklin, Methods of Mathematical Economics*, for an elementary (but long) proof.)

Then prove by contradiction: suppose f has no fixed points in B. That is, $\forall x \in B$, $x \neq f(x)$. Since x and its image $f(x)$ are distinct points in B for every x, we can carry out the following construction. For each $x \in B$, construct the line segment originating at $f(x)$ and going through x. Let $g(x)$ denote the intersection of this line segment with ∂B. This construction gives a continuous function $g : B \to \partial B$. Furthermore, notice that if $x_0 \in \partial B$, then $x_0 = g(x_0)$. Then, g gives $g(x) = x, \forall x \in \partial B$. Since there are no such functions by the fact above, we have a contradiction.

Chapter 7 Correspondence: $\Psi : X \to 2^Y$ (@ Lec 07 of ECON 204)

Definition 7.1 (Correspondence)

A **correspondence** $\Psi : X \to 2^Y$ from X to Y is a function from X to 2^Y, that is, $\Psi(x) \subseteq Y$ for every $x \in X$. (2^Y is the set of all subsets of Y) ♣

Example 7.1 Let $u : \mathbb{R}^n_+ \to \mathbb{R}$ be a continuous utility function, $y > 0$ and $p \in \mathbb{R}^n_{++}$, that is, $p_i > 0$ for each i. Define $\Psi : \mathbb{R}^n_{++} \times \mathbb{R}_{++} \to 2^{\mathbb{R}^n_+}$ by

$$\Psi(p, y) = \operatorname{argmax} u(x)$$
$$\text{s.t. } x \geq 0$$
$$p \cdot x \leq y$$

Ψ is the demand correspondence associated with the utility function u; typically $\Psi(p, y)$ is multi-valued.

7.1 Continuity of Correspondences

7.1.1 Upper/Lower Hemicontinuous

Let $X \subseteq \mathbb{E}^n$, $Y \subseteq \mathbb{E}^m$, and $\Psi : X \to 2^Y$.

Definition 7.2 (Upper Hemicontinuous)

Ψ is **upper hemicontinuous** (uhc) at $x_0 \in X$ if, for every open set V with $\Psi(x_0) \subseteq V$, there is an open set U with $x_0 \in U$ s.t.

$$x \in U \Rightarrow \Psi(x) \subseteq V$$

♣

Upper hemicontinuity reflects the requirement that Ψ doesn't "jump down/implode in the limit" at x_0. *(A set to "jump down" at the limit x_0: It should mean the set suddenly gets smaller – it "implodes in the limit" – that is, there is a sequence $x_n \to x_0$ and points $y_n \in \Psi(x_n)$ that are far from every point of $\Psi(x_0)$ as $n \to \infty$.)*

Definition 7.3 (Lower Hemicontinuous)

Ψ is **lower hemicontinuous** (lhc) at $x_0 \in X$ if, for every open set V with $\Psi(x_0) \cap V \neq \emptyset$, there is an open set U with $x_0 \in U$ s.t.

$$x \in U \Rightarrow \Psi(x) \cap V \neq \emptyset$$

♣

Lower hemicontinuity reflects the requirement that Ψ doesn't "jump up/explode in the limit" at x_0. *(A set to "jump up" at the limit x_0: It should mean that the set suddenly gets bigger – it "explodes in the limit" – that is,* *there is a sequence $x_n \to x_0$ and a point $y_0 \in \Psi(x_0)$ that is far from every point of $\Psi(x_n)$ as $n \to \infty$.)*

Definition 7.4 (Continuous Correspondence)

Ψ is **continuous** at $x_0 \in X$ if it is both **uhc** and **lhc** at x_0. ♣

Proposition 7.1

Ψ is upper hemicontinuous (respectively lower hemicontinuous, continuous) if it is uhc (respectively lhc, continuous) at every $x \in X$. ♠

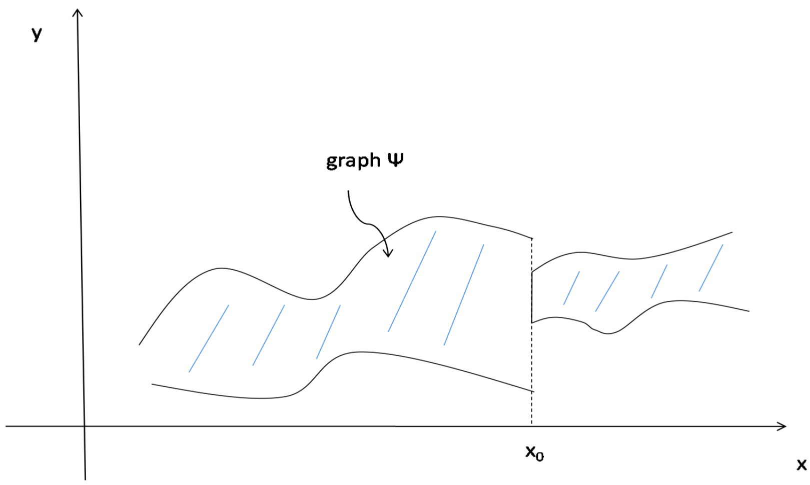

Figure 7.1: The correspondence Ψ "implodes in the limit" at x_0. Ψ is not upper hemicontinuous at x_0.

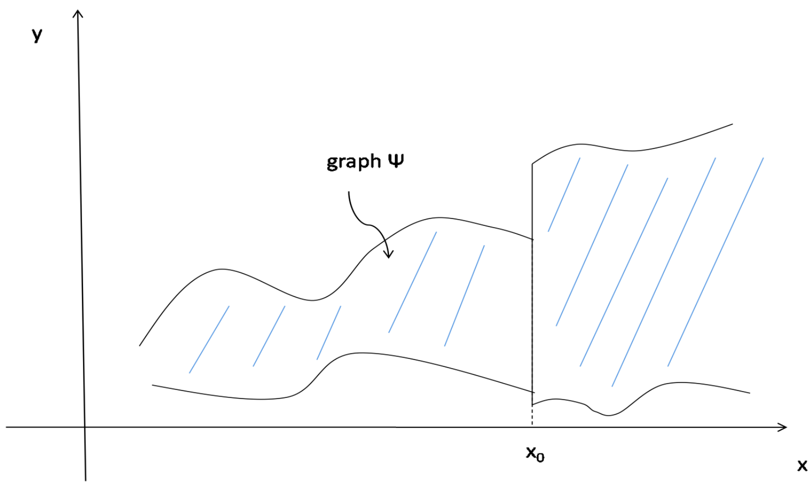

Figure 7.2: The correspondence Ψ "explodes in the limit" at x_0. Ψ is not lower hemicontinuous at x_0.

7.1.2 Theorem: $\Psi(x) = \{f(x)\}$ is uhc $\Leftrightarrow$ f is continuous

Theorem 7.1 ($\Psi(x) = \{f(x)\}$ **is uhc** $\Leftrightarrow$ f **is continuous)**

Let $X \subseteq \mathbb{E}^n$, $Y \subseteq \mathbb{E}^m$ *and* $f : X \to Y$. *Let* $\Psi : X \to 2^Y$ *be defined by* $\Psi(x) = \{f(x)\}$ *for all* $x \in X$. *Then* Ψ *is* ***uhc*** *if and only if* f *is* ***continuous***. ♡

7.1.3 Berge's Maximum Theorem: the set of maximizers is uhc with non-empty compact values

Theorem 7.2 (Berge's Maximum Theorem)

Let $X \subseteq \mathbb{R}^n$ *and* $Y \subseteq \mathbb{R}^m$. *Consider the function* $f : X \times Y \to \mathbb{R}$ *and the correspondence* $\Gamma : Y \to 2^X$. *Define* $v(y) = \max_{x \in \Gamma(y)} f(x, y)$ *and the set of maximizers*

$$\Omega(y) = \underset{x \in \Gamma(y)}{\operatorname{argmax}} f(x, y) = \{x : f(x, y) = v(y)\}$$

Suppose f *and* Γ *are continuous, and that* Γ *has non-empty compact values. Then,* v *is continuous and* Ω *is uhc with non-empty compact values.* ♡

7.2 Graph of Correspondence

An alternative notion of continuity looks instead at properties of the graph of the correspondence.

Definition 7.5 (Graph of Correspondence)

The **graph** of a correspondence $\Psi : X \to 2^Y$ is the set

$$\text{graph}\Psi = \{(x, y) \in X \times Y : y \in \Psi(x)\}$$ ♣

7.2.1 Closed Graph

By the definition of continuous function $f : \mathbb{R}^n \to \mathbb{R}$, each convergent sequence $\{(x_n, y_n)\}$ in graph f converges to a point (x, y) in graph f, that is, graph f is closed.

Definition 7.6 (Closed Graph)

Let $X \subseteq \mathbb{E}^n$, $Y \subseteq \mathbb{E}^m$. A correspondence $\Psi : X \to 2^Y$ has closed graph if its graph is a closed subset of $X \times Y$, that is, if for any sequences $\{x_n\} \subseteq X$ and $\{y_n\} \subseteq Y$ such that $x_n \to x \in X$, $y_n \to y \in Y$ and $y_n \in \Psi(x_n)$ for each n, then $y \in \Psi(x)$. ♣

Example 7.2 Consider the correspondence $\Psi(x) = \begin{cases} \{\frac{1}{x}\}, & \text{if } x \in (0, 1] \\ \{0\}, & \text{if } x = 0 \end{cases}$ ("implode in the limit")

Let $V = (-0.1, 0.1)$. Then $\Psi(0) = \{0\} \subseteq V$, but no matter how close x is to 0, $\Psi(x) = \{\frac{1}{x}\} \nsubseteq V$, so Ψ is not uhc at 0. However, note that Ψ has closed graph.

7.3 Closed-valued, Compact-valued, and Convex-valued Correspondences

Definition 7.7 (Closed-valued, Compact-valued, and Convex-valued Correspondences)

Given a correspondence $\Psi : X \to 2^Y$,

1. Ψ is **closed-valued** if $\Psi(x)$ is a closed subset of Y for all x;
2. Ψ is **compact-valued** if $\Psi(x)$ is compact for all x.
3. Ψ is **convex-valued** if $\Psi(x)$ is convex for all x.

♣

7.3.1 Closed-valued, uhc and Closed Graph

For closed-valued correspondences these concepts can be more tightly connected. A closed-valued and upper hemicontinuous correspondence must have closed graph. For a closed-valued correspondence with a compact range, upper hemicontinuity is equivalent to closed graph.

Theorem 7.3

Let $X \subseteq \mathbb{E}^n$, $Y \subseteq \mathbb{E}^m$, and $\Psi : X \to 2^Y$.

1. *Ψ is **closed-valued** and **uhc** $\Rightarrow$ Ψ has **closed graph**.*
2. *Ψ is **closed-valued** and **uhc** $\Leftarrow$ Ψ has **closed graph**. (If Y is **compact**)*

♡

Theorem 7.4

*Let $X \subseteq \mathbb{E}^n$, $Y \subseteq \mathbb{E}^m$, and $\Psi : X \to 2^Y$. If Ψ has **closed graph** and there is an **open set** W with $x_0 \in W$ and a **compact set** Z such that $x \in W \cap X \Rightarrow \Psi(x) \subseteq Z$, then Ψ is **uhc** at x_0.*

♡

7.3.2 Theorem: compact-valued, uhs correspondence of compact set is compact

Theorem 7.5

Let X be a compact set and $\Psi : X \to 2^X$ be a non-empty, compact-valued upper-hemicontinuous correspondence. If $C \subseteq X$ is compact, then $\Psi(C)$ is compact.

♡

Proof 7.1

Given the compact-valued Ψ, we can have an open cover of $\Psi(C)$, $\{U_\lambda : \lambda \in \Lambda\}$. So $\forall x \in C$, there exists $U_{l(x)}, l(x) \in \Lambda$ such that $U_{l(x)}$ is an open cover of $\Psi(x)$.

Consider a $c \in C$. Since Ψ is uhs and $\Psi(c) \subseteq U_{l(c)}$, there exists open set V_c s.t. $c \in V_c$ and $\Psi(x) \subseteq U_{l(c)}, \forall x \in V_c \cap C$.

$\{V_c : c \in C\}$ is an open cover of C. Because C is compact, there is a finite subcover $\{V_{c_i} : i = 1, ..., m\}, m \in \mathbb{N}$, where $\{c_i : i = 1, ..., m\} \subseteq C$.

Because $\Psi(x) \subseteq U_{l(c_i)}, \forall x \in V_{c_i} \cap C$ and $\{V_{c_i} : i = 1, ..., m\}, m \in \mathbb{N}$ is a open cover for C, we can infer $\{U_{l(c_i)} : i = 1, ..., m\}$ is a finite subcover of $\{U_{l(c)} : c \in C\}$ for $\Psi(C)$. Hence, $\Psi(C)$ is compact.

7.4 Fixed Points for Correspondences (@ Lec 13 of ECON 204)

7.4.1 Definition

Definition 7.8 (Fixed Points for Correspondences)

Let X be nonempty and $\psi : X \to 2^X$ be a correspondence. A point $x^* \in X$ is a fixed point of ψ if $x^* \in \psi(x^*)$.

Note *We only need x^* to be in $\psi(x^*)$, not $\{x^*\} = \psi(x^*)$. That is, ψ need not be single-valued at x^*. So x^* can be a fixed point of ψ but there may be other elements of $\psi(x^*)$ different from x^*.*

7.4.2 Kakutani's Fixed Point Theorem: uhs, compact, convex values correspondence has a fixed point over compact convex set

Theorem 7.6 (Kakutani's Fixed Point Theorem)

*Let $X \subseteq \mathbb{R}^n$ be a non-empty, **compact**, **convex** set and $\psi : X \to 2^X$ be an **upper hemi-continuous** correspondence with non-empty, **compact**, **convex** values. Then ψ has a fixed point in X.*

7.4.3 Theorem: $\exists$ compact set $C = \cap_{i=0}^{\infty} \Psi^i(X)$ s.t. $\Psi(C) = C$

Theorem 7.7

Let (X, d) be a compact metric space and let $\Psi(x) : X \to 2^X$ be a upper-hemicontinuous, compact-valued correspondence, such that $\Psi(x)$ is non-empty for every $x \in X$. There exists a compact non-empty subset $C \subseteq X$, such that $\Psi(C) \equiv \cup_{x \in C} \Psi(x) = C$.

Proof 7.2

Let's construct a sequence $\{C_n\}$ such that $C_0 = X$, $C_1 = \Psi(C_0)$, ..., $C_n = \Psi(C_{n-1})$, ... We claim that $C = \cap_{i=0}^{\infty} C_i$ is a non-empty compact set and satisfies $\Psi(C) = C$.

1. *Because we can infer $\Psi(X_1) \subseteq \Psi(X_2)$ if $X_1 \subseteq X_2$, $X = C_0 \supseteq C_1 \Rightarrow C_1 = \Psi(C_0) \supseteq C_2 = \Psi(C_1)$,...., so $C_0 \supseteq C_1 \supseteq \cdots C_n \supseteq \cdots$. Hence, C is not empty.*

2. *Because X is compact, by the theorem 7.5, we can infer C_n is compact for all n. Then, C_n is closed for all n, so C is closed. Because C is a closed set of compact set X, C is compact.*
3. $C \subseteq C_n, \forall n \Rightarrow \Psi(C) \subseteq \Psi(C_n), \forall n \Rightarrow \Psi(C) \subseteq C$
4. *Assume $C \subseteq \Psi(C)$ doesn't hold, that is $\exists y \in C$ s.t. $y \notin \Psi(C)$. Because $y \in C$ and $C_0 \supseteq C_1 \supseteq \cdots C_n \supseteq \cdots$, there exists $k \in C_n$ for all n s.t. $y \in \Psi(k)$. $k \in \cap_{i=1}^{\infty} C_i = C$, so $\Psi(k) \subseteq \Psi(C)$, which contradicts to $y \notin \Psi(C)$. Hence, $C \subseteq \Psi(C)$.*

All in all the claim "$C = \cap_{i=0}^{\infty} C_i$ is a non-empty compact set and satisfies $\Psi(C) = C$" is proved.

Chapter 8 Bayesian Persuasion: Extreme Points and Majorization

Based on

- Kleiner, A., Moldovanu, B., & Strack, P. (2021). Extreme points and majorization: Economic applications. *Econometrica*, 89(4), 1557-1593.
-

8.1 Extreme Points

8.1.1 Extreme Points of Convex Set

Definition 8.1 (Extreme Points)

An **extreme point** of a convex set A is a point $x \in A$ that cannot be represented as a convex combination of points in A.

♣

8.1.2 Krein-Milman Theorem: Existence of Extreme Points

Theorem 8.1 (Krein-Milman Theorem)

*Every non-empty **compact convex** subset of a Hausdorff locally convex topological vector space (for example, a normed space) is the closed, convex hull of its extreme points.*

In particular, this set has extreme points.

♡

8.1.3 Bauer's Maximum Principle: Usefulness of Extreme Points for Optimization

Theorem 8.2 (Bauer's Maximum Principle)

*Any function that is **convex and continuous**, and defined on a set that is **convex and compact**, attains its maximum at some extreme point of that set.*

♡

8.2 Majorization

8.2.1 Majorization and Weak Majorization

Definition 8.2 (Majorization of Non-decreasing Functions)

Consider right-continuous functions that map the unit interval $[0,1]$ into the real numbers. For two non-decreasing functions $f, g \in L^1$, we say that f **majorizes** g, denoted by $g \prec f$, if the following two conditions hold:

$$\int_x^1 g(s)ds \leq \int_x^1 f(s)ds, \forall x \in [0,1] \qquad \text{(Condition 1)}$$

$$\int_0^1 g(s)ds = \int_0^1 f(s)ds \qquad \text{(Condition 2)}$$

♣

Definition 8.3 (Weak Majorization)

f **weakly majorizes** g, denoted by $g \prec_w f$, if Condition 1 holds (not necessarily Condition 2). ♣

8.2.2 How to work for non-monotonic functions? – Non-Decreasing Rearrangement

Note How this work with non-monotonic functions?

Suppose f, g are non-monotonic, we compare their non-decreasing rearrangements f^*, g^*.

Definition 8.4 (Rearrangement)

Given a function f, let $m(x)$ denote the Lebesgue measure of the set $\{s \in [0,1] : f(s) \leq x\}$, that is $m(x) = \int_{s\in\{s\in[0,1]:f(s)\leq x\}} 1ds$ (the "length" of the set). The non-decreasing rearrangement of f, f^*, is defined by

$$f^*(t) = \inf\{x \in \mathbb{R} : m(x) \geq t\},\ t \in [0,1]$$

♣

8.2.3 Theorem: F majorizes $G \Leftrightarrow G$ is a mean-preserving spread of F

Based on

- Shaked, M., & Shanthikumar, J. G. (2007). *Stochastic orders*. New York, NY: Springer New York.

Definition 8.5 (Generalized Inverse)

Suppose G is defined on the interval $[0,1]$, we can define the **generalized inverse**

$$G^{-1}(x) = \sup\{s : G(s) \leq x\}, x \in [0,1]$$

♣

Let X_F and X_G be now random variables with distributions F and G, defined on the interval $[0,1]$.

Theorem 8.3 (Shaked & Shanthikumar (2007), Section 3.A)

$$G \prec F \Leftrightarrow F^{-1} \prec G^{-1} \Leftrightarrow X_G \leq_{ssd} X_F \text{ and } \mathbb{E}[X_G] = \mathbb{E}[X_F]$$

where $\leq_{ssd}$ denotes the standard second-order stochastic dominance.

Based on Theorem 3.1 and the Condition 2 of Majorization, we can conclude

Corollary 8.1 (Majorization $\Leftrightarrow$ Mean-preserving Contraction)

F majorizes $G \Leftrightarrow F$ is a mean-preserving contraction of G (G is a mean-preserving spread of F)

That is, we can construct random variables X_F, X_G, jointly distributed on some probability space, such that $X_F \sim F$, $X_G \sim G$ and such that $X_F = \mathbb{E}[X_G \mid X_F]$.

8.3 Capture Extreme Points in Economic Applications

Let L^1 denote the real-valued and integrable functions defined on $[0, 1]$.

In this section, we focus on **non-decreasing (weakly increasing) functions**, for example, a cumulative distribution function in Bayesian persuasion, or an incentive-compatible allocation in mechanism design.

8.3.1 Definitions of $\mathcal{MPS}(f)$, $\mathcal{MPS}_w(f)$, $\mathcal{MPC}(f)$

Based on Corollary 8.1, we can define following sets

Definition 8.6

1. The set of non-decreasing functions that are majorized by f is denoted by
$$\begin{aligned}\mathcal{MPS}(f) &= \text{MPS}(f) \cap \{g \in L^1 \mid g \text{ is non-decreasing}\}\\ &= \{g \in L^1 \mid g \text{ is non-decreasing and } g \prec f\}\end{aligned}$$
2. The set of non-negative, non-decreasing functions that are weakly majorized by f is denoted by
$$\mathcal{MPS}_w(f) = \{g \in L^1 \mid g \text{ is non-negative, non-decreasing and } g \preceq f\}$$
3. The set of non-decreasing functions that majorize f and satisfy $f(0) \leq g \leq f(1)$ is denoted by
$$\begin{aligned}\mathcal{MPC}(f) &= \text{MPC}(f) \cap \{g \in L^1 \mid g \text{ is non-decreasing and } f(0) \leq g \leq f(1)\}\\ &= \{g \in L^1 \mid g \text{ is non-decreasing and } g \succ f \text{ and } f(0) \leq g \leq f(1)\}\end{aligned}$$
where $f(0) \leq g \leq f(1)$ is used to ensure compactness.

8.3.2 Proposition: $\mathcal{MPS}(f)$, $\mathcal{MPS}_w(f)$, $\mathcal{MPC}(f)$ have extreme points and any element is a combination of extreme points

Following two propositions are the Proposition 1 of the Kleiner et al. (2021).

Proposition 8.1 (Non-decreasing $f \Rightarrow \mathcal{MPS}(f)$, $\mathcal{MPS}_w(f)$, and $\mathcal{MPC}(f)$ have extreme points)

Suppose $f \in L^1$ is non-decreasing. Then $\mathcal{MPS}(f)$, $\mathcal{MPS}_w(f)$, and $\mathcal{MPC}(f)$ are convex and compact in the norm topology $\Rightarrow$ (by Krein-Milman Theorem 8.1) they all have non-empty set of extreme points.

Note *We use* ext$A$ *to denote the set of extreme points of set A.*

Proposition 8.2 (Non-decreasing $f \Rightarrow$ any distribution is a combination of extreme points)

Suppose $f \in L^1$ is non-decreasing. For any $g \in \mathcal{MPS}(f)$, $\exists$ a probability measure λ_g over ext$\mathcal{MPS}(f)$ *such that*

$$g = \int_{\text{ext}\mathcal{MPS}(f)} h \, d\lambda_g(h)$$

(also hold for any $g \in \mathcal{MPS}_w(f)$ and $g \in \mathcal{MPC}(f)$).

8.3.3 Extreme Points in $\mathcal{MPS}(f)$

Theorem 8.4 (Form of Extreme Points in MPS(f): Kleiner et al. (2021), Theorem 1)

*Let f be non-decreasing. Then g is an **extreme point** in $\mathcal{MPS}(f)$ if and only if there exists a collection of disjoint intervals $\{[\underline{x}_i, \overline{x}_i)\}_{i \in I}$ such that*

$$g(x) = \begin{cases} f(x), & \text{if } x \notin \cup_{i \in I}[\underline{x}_i, \overline{x}_i) \\ \frac{\int_{\underline{x}_i}^{\overline{x}_i} f(s)ds}{\overline{x}_i - \underline{x}_i}, & \text{if } x \in [\underline{x}_i, \overline{x}_i) \end{cases}$$

g is an extreme point of $\mathcal{MPS}(f)$ implies either that $g(x) = f(x)$ or that g is constant at x.

Definition 8.7 (Exposed Element)

An element x of a convex set A is **exposed** if there exists a continuous linear functional that attains its maximum on A uniquely at x.

Note *Every exposed point is extreme, but the converse is not true in general.*

Corollary 8.2 (Kleiner et al. (2021), Corollary 1)

Every extreme point of $\mathcal{MPS}(f)$ is exposed.

8.3.4 Extreme Points in $\mathcal{MPS}_w(f)$

For a set $A \subseteq [0,1]$, we use $\mathbf{1}_A(x)$ denote the indicator function of set A: it equals to 1 if $x \in A$ and 0 otherwise.

Corollary 8.3 (Kleiner et al. (2021), Corollary 2)

Suppose that f is non-decreasing and non-negative. A function g is an extreme point of $\mathcal{MPS}_w(f)$ if and only if there is $\theta \in [0,1]$ such that g is an extreme point of $\mathcal{MPS}(f)$ and $g(x) = 0, \forall x \in [0,\theta)$. ♡

8.3.5 Extreme Points in $\mathcal{MPC}(f)$

Theorem 8.5 (Kleiner et al. (2021), Theorem 2)

Let f be non-decreasing and continuous. Then $g \in \mathcal{MPC}(f)$ is an extreme point of $\mathcal{MPC}(f)$ if and only if there exists a collection of intervals $[\underline{x}_i, \bar{x}_i)$, (potentially empty) sub-intervals $\left[\underline{y}_i, \bar{y}_i\right) \subseteq [\underline{x}_i, \bar{x}_i)$, and numbers v_i indexed by $i \in I$ such that for all $x \in [0,1]$,

$$g(x) = \begin{cases} f(x) & \text{if } x \notin \bigcup_{i \in I} [\underline{x}_i, \bar{x}_i) \\ f(\underline{x}_i) & \text{if } x \in \left[\underline{x}_i, \underline{y}_i\right) \\ v_i & \text{if } x \in \left[\underline{y}_i, \bar{y}_i\right) \\ f(\bar{x}_i) & \text{if } x \in [\bar{y}_i, \bar{x}_i) \end{cases} \tag{8.1}$$

Moreover, a function g as defined in (8.1) is in $\mathcal{MPC}(f)$ if the following three conditions are satisfied:

$$\left(\bar{y}_i - \underline{y}_i\right) v_i = \int_{\underline{x}_i}^{\bar{x}_i} f(s)\mathrm{d}s - f(\underline{x}_i)\left(\underline{y}_i - \underline{x}_i\right) - f(\bar{x}_i)(\bar{x}_i - \bar{y}_i) \tag{8.2}$$

$$f(\underline{x}_i)(\bar{y}_i - \underline{x}_i) + f(\bar{x}_i)(\bar{x}_i - \bar{y}_i) \leq \int_{\underline{x}_i}^{\bar{x}_i} f(s)\mathrm{d}s \leq f(\underline{x}_i)\left(\underline{y}_i - \underline{x}_i\right) + f(\bar{x}_i)\left(\bar{x}_i - \underline{y}_i\right) \tag{8.3}$$

If $v_i \in \left(f\left(\underline{y}_i\right), f(\bar{y}_i)\right)$, then for an arbitrary point m_i satisfying $f(m_i) = v_i$ it must hold that

$$\int_{m_i}^{\bar{x}_i} f(s)\mathrm{d}s \leq v_i(\bar{y}_i - m_i) + f(\bar{x}_i)(\bar{x}_i - \bar{y}_i) \tag{8.4}$$

♡

Condition (8.2) in the theorem ensures that g and f have the same integrals for each sub-interval $[\underline{x}_i, \bar{x}_i)$, analogously to the condition imposed in Theorem 8.3.3. Condition (8.3) ensures that $v_i \in (f(\underline{x}_i), f(\bar{x}_i))$, ensuring that g is non-decreasing. If f crosses g in the interval $\left[\underline{y}_i, \bar{y}_i\right]$, then there is $m_i \in \left[\underline{y}_i, \bar{y}_i\right]$ such that $f(m_i) = v_i$. In this case, Condition (8.4) ensures that $\int_s^{\bar{x}_i} f(t)\mathrm{d}t \leq \int_s^{\bar{x}_i} g(t)\mathrm{d}t$ for all $s \in [\underline{x}_i, \bar{x}_i)$ and thus that $f \prec g$. If $v_i \notin \left(f\left(\underline{y}_i\right), f(\bar{y}_i)\right)$, Condition (8.3) is enough to ensure that $f \prec g$ and thus Condition (8.4) is not necessary.

Chapter 9 Bayesian Persuasion: Bi-Pooling

Based on

- ★ Arieli, I., Babichenko, Y., Smorodinsky, R., & Yamashita, T. (2023). Optimal persuasion via bi-pooling. *Theoretical Economics*, 18(1), 15-36.
- Gentzkow, Matthew and Emir Kamenica (2016), "A Rothschild-Stiglitz approach to Bayesian persuasion." *American Economic Review*, 106, 597-601.
- Kolotilin, Anton (2018), "Optimal information disclosure: A linear programming ap-proach." *Theoretical Economics*, 13, 607-635.

9.1 Persuation Model

Consider a persuation model where the state space is the interval $[0, 1]$ with a common prior $F \in \Delta([0, 1])$ that has full support (i.e., $[0, 1]$ is the smallest closed set that has probability one). The sender knows the realized state and the receiver is uninformed.

1. Singaling: Prior to the realization of the state, the sender commits to a **signaling policy**

$$\pi : [0, 1] \to \Delta(S)$$

   where S is an arbitrary measurable space. Once the state $\omega \in [0, 1]$ is realized, the sender sends a signal $s \in S$ to the receiver based on the committed signaling policy, i.e., $s \sim \pi(\omega)$. Without loss of generality, we may assume that $S = [0, 1]$, and that the posterior mean of the state, given signal s, is s itself.

   Hence, the distribution of the posterior mean s given the signal policy π, denoted by $F_\pi \in \Delta([0, 1])$ is a *mean-preserving contraction* of F (i.e., $\exists \varepsilon_\omega \in \Delta([0, 1])$ such that $\omega = s + \varepsilon_\omega$ for all $\omega \in F$ and $s \in F_\pi$). It is also easy to note that for any $G \in \text{MPC}(F)$, there exists a signaling policy π (may not be unique) that makes $F_\pi = G$ (e.g., Gentzkow and Kamenica(2016), Kolotilin (2018)).

2. Persuation problem: The sender's indirect utility is denoted by $u : [0, 1] \to \mathbb{R}$, where $u(x)$ is the sender's expected utility in case the receiver's posterior mean is x. u is assume to be upper semicontinuous. (F, u) is referred as a **persuation problem**. The sender's problem takes the form:

$$\max_{G \in \text{MPC}(F)} \mathbb{E}_{x \sim G}[u(x)]$$

9.2 Bi-Pooling

9.2.1 Bi-pooling Distribution

Note *For a distribution $H \in \Delta([0,1])$ and a measurable set $C \subseteq [0,1]$ we denote by $H|_C$ the distribution of $h \sim H$ conditional on the event that $h \in C$.*

Definition 9.1 (Bi-pooling Distribution (Arieli et al. (2023), Definition 1))

A distribution $G \in \text{MPC}(F)$ is called a **bi-pooling distribution** (with respect to F) if there exists a collection of pairwise disjoint open intervals $\{(\underline{y}_i, \overline{y}_i)\}_{i \in A}$ such that

- For every $i \in A$,

$$G((\underline{y}_i, \overline{y}_i)) = F((\underline{y}_i, \overline{y}_i))$$

  where $G((\underline{y}_i, \overline{y}_i)) = G(\overline{y}_i) - G(\underline{y}_i) = \int_{\underline{y}_i}^{\overline{y}_i} g(x)dx$, $F((\underline{y}_i, \overline{y}_i)) = F(\overline{y}_i) - F(\underline{y}_i) = \int_{\underline{y}_i}^{\overline{y}_i} f(x)dx$.

- The remaining intervals are the same:

$$G|_{[0,1] \setminus \cup_{i \in A}(\underline{y}_i, \overline{y}_i)} = F|_{[0,1] \setminus \cup_{i \in A}(\underline{y}_i, \overline{y}_i)}$$

- For every $i \in A$,

$$|\text{supp}(G|_{(\underline{y}_i, \overline{y}_i)})| \leq 2$$

  which means there are at most two different values of G over $(\underline{y}_i, \overline{y}_i)$. If $|\text{supp}(G|_{(\underline{y}_i, \overline{y}_i)})| = 2$, we call $(\underline{y}_i, \overline{y}_i)$ a **bi-pooling interval**; If $|\text{supp}(G|_{(\underline{y}_i, \overline{y}_i)})| = 1$, we call $(\underline{y}_i, \overline{y}_i)$ a **pooling interval**. In the case where all intervals are pooling intervals, we say that G is a **pooling distribution** (with respect to F).

♣

Example 9.1 Consider the persuasion problem (F, u), where $F = U[0,1]$ is the uniform distribution over $[0,1]$ and $u : [0,1] \to \mathbb{R}$ is an arbitrary function satisfying $u(\frac{1}{3}) = u(\frac{2}{3}) = 0$ and $u(x) < 0, \forall x \notin \{\frac{1}{3}, \frac{2}{3}\}$.
Consider using a binary signal space $S = \{s_1, s_2\}$, where s_1 is sent with probability 1 over the interval $(\frac{1}{12}, \frac{7}{12})$ and s_2 is sent with probability 1 over the interval $[0, \frac{1}{12}] \cup [\frac{7}{12}, 1]$. This policy is a bi-pooling policy for the singleton collection $\{[0,1]\}$.

9.3 Applying Bi-pooling Distributions to Persuasion Problems

9.3.1 It works for all

Theorem 9.1 (Arieli et al. (2023), Theorem 1)

Every persuasion problem (F, u) *admits an optimal bi-pooling distribution.* ♡

Proposition 9.1 (Arieli et al. (2023), Proposition 1)

The set of extreme points of $\mathrm{MPC}(F)$ *is precisely the set of bi-pooling distributions.* ♠

Theorem 9.2 (Arieli et al. (2023), Theorem 2)

For every bi-pooling distribution $G \in \mathrm{MPC}(F)$ *there exists a continuous utility function* u *for which* G *is the unique optimal solution of* $\max_{G \in \mathrm{MPC}(F)} \mathbb{E}_{x \sim G}[u(x)]$. *That is, every extreme point of* $\mathrm{MPC}(F)$ *is exposed.* ♡

9.3.2 How it works

Definition 9.2 (Bi-pooling Policy (Arieli et al. (2023), Definition 3))

A signaling policy π is called a **bi-pooling policy** if there exists a collection of pairwise disjoint intervals $\{(\underline{y}_i, \overline{y}_i)\}_{i \in A}$ such that

- for every state $\omega \in (\underline{y}_i, \overline{y}_i)$ we have $\mathrm{supp}(\pi(\omega)) \subseteq \{\underline{z}_i, \overline{z}_i\}$ (either $\pi(\omega) = \overline{z}_i$ or $\pi(\omega) = \underline{z}_i$) for some $\underline{z}_i \le \overline{z}_i$ and $\underline{z}_i, \overline{z}_i \in [\underline{y}_i, \overline{y}_i]$;
- for every $\omega \notin \cup_{i \in A} (\underline{y}_i, \overline{y}_i)$, the policy sends the signal $\pi(\omega) = \omega$ (i.e., it reveals the state).

In the case where $\underline{z}_i = \overline{z}_i$ for all $i \in A$, we refer to π as a **pooling policy**. ♣

Definition 9.3 (Monotonic Signaling Policy (Arieli et al. (2023), Definition 4))

A (possibly mixed) signaling policy, $\pi : [0, 1] \to \Delta([0, 1])$, is **monotonic** if

$\pi(x)$ <u>first-order stochastically dominates</u> $\pi(y)$ for every $x \ge y$. ♣

Proposition 9.2 (Arieli et al. (2023), Proposition 2)

Every persuasion problem admits an optimal (mixed) monotonic signaling policy. ♠

Lemma 9.1 (Arieli et al. (2023), Lemma 3)

A persuasion problem (F, u) *admits an optimal pure monotonic signaling policy* <u>*if and only if*</u> *it admits an optimal pooling policy.* ♡

Definition 9.4 (Double-Interval Nested Structure)

A pure signaling policy: for each bi-pooling interval $(\underline{y}_i, \overline{y}_i)$, we can find a sub-interval $(\underline{w}_i, \overline{w}_i) \subseteq (\underline{y}_i, \overline{y}_i)$ such that π is constant over the interval $(\underline{w}_i, \overline{w}_i)$ as well as over its complement $(\underline{y}_i, \overline{y}_i) \backslash (\underline{w}_i, \overline{w}_i)$. ♣

Corollary 9.1 (Arieli et al. (2023), Corollary 2)

Every persuasion problem (F, u) has an optimal bi-pooling policy that has a double-interval nested structure. ♡