Econ 701A Slides
Decision Theory Foundations (MWG 1)

Steven A. Matthews   University of Pennsylvania

August 30, 2017

Introduction

- **Decision theory (DT)** is the study of individual decision making. (Game theory is its extension to groups of individuals.)
- Much of DT is concerned with formulating a notion of "rational" decision making, and finding conditions under which observed choices are consistent with rationality.
- DT is the basis of the **positive** predictions (**comparative statics**) and **normative** welfare (policy) prescriptions in consumer and producer theory.
- Two approaches to DT:
  - **Choice behavior** as the primitive
  - **Preference** as the primitive

Alternatives

- Both approaches start with a universal set of **alternatives**, X.
- Each $x \in X$ is supposed to be a complete description of everything the decision maker (DM) might "care about" and which could be different in some other alternative $y \in X$.

Examples

- $X = \{\text{Clinton, Trump}\}$
- $X = \{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 = \text{bread in g/day}, x_2 = \text{butter in lbs/day}\}$
- $X = \{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 = \text{\$ today}, x_2 = \text{\$ tomorrow}\}$
- $X = \{(x_1, x_2, \ldots) \in \mathbb{R}^\infty_+ : x_t = \text{\$ in period } t\}$
- $X = \{(x_1, \ldots, x_S) \in \mathbb{R}^{SL}_+ : x_s = \text{consumption bundle of } L \text{ goods in the state of nature } s = 1, \ldots, S\}$

Describing Alternatives: Framing

Example (Tversky and Kahneman)

Flu outbreak in town of 600 people. The DM must choose one of two treatment strategies:

- **Treatment A**: 400 will die.
  Treatment B: with probability $1/3$ nobody dies; with probability $2/3$ they all die.
  - DM chooses B.
- Now the DM must instead choose between two other treatment strategies:
  Treatment C: 200 people will be saved.
  Treatment D: with probability $1/3$ all are saved; with probability $2/3$ none are saved.
  - DM chooses C.

The Preference-Based Approach

- The primitive datum in this approach is a binary ("preference") relation $\succsim$ on X.

  Definition. A **binary relation** $\succsim$ on X is a subset of X^2.

  We usually (but not always!) write $x \succsim y$ rather than $\succsim \subseteq X^2$, but the meaning is the same.
- The interpretation of $x \succsim y$ is that "x is weakly preferred to y" or "x is at least as good as y". These words convey that $\succsim$ is meant to describe how the DM "feels" about different alternatives.
- In the preference-based approach, choice behavior is derived from preferences. Given $\succsim$ and a feasible set $B \subseteq X$, the prediction is that the DM will choose an alternative in the set

$$C^*(B, \succsim) := \{x \in B : x \succsim y \ \forall y \in B\}.$$

Stable Tastes and their Separation from Feasibility

Two related implicit assumptions of the preference-based approach $C^*(B, \succsim)$:

1. The DM's preferences are "stable" over time.
2. The DM's preferences do not depend on the currently feasible set of alternatives.

Example 1 (Luce and Raiffa).

- First the DM is offered menu $A = \{\text{steak tartare, chicken}\}$, from which he chooses chicken.
- Later, the DM is offered $B = \{\text{steak tartare, chicken, frog legs}\}$, from which he chooses steak tartare.

Example 2. Addiction ("changing tastes").

Rationality in the Preference-Based Approach

In the preference-based approach, the assumption that the DM is "rational" is embodied by two assumptions ("axioms") about the preference relation $\succsim$, which are that it is

1. **complete** (for all $x, y \in X$, $x \succsim y$ or $y \succsim x$), and
2. **transitive** (for all $x, y, z \in X$, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$).

If $\succsim$ is complete and transitive, MWG call it a **rational preference relation**. (But be aware that many texts define a "preference relation" to be any binary relation that is complete and transitive.)

Note that if $\succsim$ is complete, then it is **reflexive**:

$$x \succsim x \text{ for all } x \in X.$$

How Plausible Is it that Preferences Are Rational?

- One view: who cares? DT is the theory of perfectly rational decision making, meant to be a benchmark towards which imperfectly rational real-world people should strive. (normative view of DT)
- Another view: completeness is merely a technical requirement, but transitivity is violated in at least two plausible scenarios. We need more definitions and propositions to discuss them.

Strict Preference and Indifference Relations

Given a binary relation $\succsim$ on X, define two others:

- **strict preference:** $x \succ y$ iff $x \succsim y$ and not $y \succsim x$,
- **indifference:** $x \sim y$ iff $x \succsim y$ and $y \succsim x$.

Lemma

If $\succsim$ is complete and transitive, then both $\succ$ and $\sim$ are transitive. Furthermore, for any $x, y, z \in X$,

$$x \succ y \text{ and } y \succsim z \;\Rightarrow\; x \succ z,$$

$$x \succsim y \text{ and } y \succ z \;\Rightarrow\; x \succ z.$$

Aggregation of Preferences

Each family member $i \in F = \{Dad, Mom, Child\}$ has a rational preference relation $\succsim_i$. The family's preferences $\succsim$ are defined by majority rule:

$$x \succsim y \text{ iff } |\{i : x \succsim_i y\}| \geq 2.$$

Then $\succ$, and hence $\succsim$, is intransitive if the profile of individual preferences is that of the **Condorcet paradox**:

$$x \succ_D y \succ_D z, \quad y \succ_M z \succ_M x, \quad z \succ_C x \succ_C y.$$

Or: let an alternative be a cup of coffee described by a triple $($*warmth*, *flavor*, *caffeine*$) \in \mathbb{R}^3$. The DM has three **internal selves**, one caring only about warmth, another only about flavor, and another only about caffeine. The DM's preference between any two cups is determined by majority rule among the selves. The DM's preferences are then intransitive: consider $x = (2, 0, 1)$, $y = (1, 2, 0)$, and $z = (0, 1, 2)$.

Plausible Scenario 2 for Intransitivity

Imperceptible Differences

- Consider a DM who likes sugar in her coffee.
- Let x_i = a cup of coffee with i grains of sugar.
- The DM likes sugar in coffee: thus, $x_{50} \succ x_0$.
- But, since she cannot taste one grain of sugar, $x_i \sim x_{i+1}$.

Hence, $\sim$ is intransitive

$\Rightarrow \succsim$ is intransitive.

Choice-Based Approach

- The primitive datum in this approach is a *choice structure*, $\langle \mathfrak{B}, C \rangle$, where

$$\mathfrak{B} \subseteq 2^X \backslash \varnothing \text{ is the set of possible feasible ('budget') sets,}$$
$$C : \mathfrak{B} \to 2^X \text{ is the } \textbf{choice correspondence,}$$

where by definition (ours and MWG's),

$$\varnothing \neq C(B) \subseteq B \text{ for all } B \in \mathfrak{B}.$$

- The advantage of this approach is that it is based on observable, measurable objects (feasible sets, choice behavior), as opposed to unobservable psychological objects (preferences).

Rationality in the Choice-Based Approach

In the choice-based approach, the assumption that the DM is "rational" is embodied by the "Weak Axiom of Revealed Preference." The first step is to extract a preference relation from a choice structure.

Definition

Given $\langle \mathfrak{B}, C \rangle$, *the associated* **revealed preference relation,** $\succsim^*$, *is defined by the following: for all* $x, y \in X$,

$$x \succsim^* y \text{ iff } x = y \text{ or } B \in \mathfrak{B} \text{ exists such that } y \in B \text{ and } x \in C(B).$$

(This definition differs from MWG's by assuming reflexivity – MWG does the same thing by assuming $\mathfrak{B}$ contains all singleton sets.)

Observation. $\succsim^*$ is complete if $\mathfrak{B}$ contains all size-two subsets.

The Weak Axiom of Revealed Preference

Definition (WARP)

A choice structure $\langle \mathfrak{B}, C \rangle$ *satisfies the* **Weak Axiom of Revealed Preference** *if whenever* x *is (has been) revealed preferred to* y*, and in some instance* x *is available and* y *is chosen, then* x *is also then chosen:* $\langle \mathfrak{B}, C \rangle$ *satisfies WARP iff for all* $B \in \mathfrak{B}$ *and* $x, y \in B$*,*

$$y \in C(B) \text{ and } x \succsim^* y \;\Rightarrow\; x \in C(B).$$

Connecting the Preference and Choice-Based Approaches

The preference-based approach is easier: it enables our mathematical optimization tools. But the choice-based approach is more soundly rooted on observable behavior.

How compatible are the two approaches?

Definition

$\langle \mathfrak{B}, C \rangle$ is **rationalized** *by a binary relation $\succsim$ if*

$$C(B) = C^*(B, \succsim) \text{ for all } B \in \mathfrak{B}.$$

We seek conditions under which a choice structure $\langle \mathfrak{B}, C \rangle$ is rationalized by a **rational** preference relation, i.e., by a complete and transitive binary relation.

The Revealed Preference Theorem

Now we obtain the result that connects the preference and choice-based approaches to decision theory. The following is equivalent to MWG Propositions 1.D.1 and 1.D.2.

Theorem

Let $\langle \mathfrak{B}, C \rangle$ be a choice structure.

1. *If $\langle \mathfrak{B}, C \rangle$ is rationalized by a transitive binary relation $\succsim$, then it satisfies WARP.*
2. *Conversely, if $\langle \mathfrak{B}, C \rangle$ satisfies WARP and $\mathfrak{B}$ contains all subsets of X of size three, then $\succsim^*$ is transitive and rationalizes $\langle \mathfrak{B}, C \rangle$. If in addition $\mathfrak{B}$ contains all subsets of X of size two, then $\succsim^*$ is also complete and is the only reflexive binary relation that rationalizes $\langle \mathfrak{B}, C \rangle$.*

Proof of Part 1

Proof that transitive rationalizability implies WARP:

- Suppose $x \succsim^* y$, $\{x, y\} \subseteq B \in \mathfrak{B}$, and $y \in C(B)$.
- Then $y \in C^*(B, \succsim)$, and so $y \succsim z$ for all $z \in B$.
- Since $x \succsim^* y$, $B' \in \mathfrak{B}$ exists such that $\{x, y\} \subseteq B'$ and $x \in C(B')$.
- Hence $x \in C^*(B', \succsim)$, and so $x \succsim y$.
- The transitivity of $\succsim$ now implies $x \succsim z$ for all $z \in B$.
- Thus, $x \in C^*(B, \succsim) = C(B)$, which proves that WARP holds. $\blacksquare$

Proof of Part 2 of the Theorem

We first show that $\succsim^*$ is transitive:

- Assume $x \succsim^* y$ and $y \succsim^* z$, but **not** $x \succsim^* z$.
- By assumption, $\{x, y, z\} \in \mathfrak{B}$.
- Because $x \succsim^* z$ does not hold, $x \notin C(\{x, y, z\})$.
- So WARP and $x \succsim^* y$ imply $y \notin C(\{x, y, z\})$.
- Now WARP and $y \succsim^* z$ imply $z \notin C(\{x, y, z\})$.
- So $C(\{x, y, z\}) = \varnothing$, contrary to C being a choice correspondence.
- Therefore, $\succsim^*$ is transitive. $\blacksquare$

Proof of Part 2 of the Theorem (con't)

We now show that $\succsim^*$ rationalizes $\langle \mathfrak{B}, C \rangle$:

- If $x \in C(B)$, then by definition $x \succsim^* y$ for all $y \in B$. Hence,

$$C(B) \subseteq C^*(B, \succsim^*) \text{ for all } B \in \mathfrak{B}.$$

- Now let $B \in \mathfrak{B}$ and $x \in C^*(B, \succsim^*)$.
- Then $x \succsim^* y$ for all $y \in B$.
- Since $C(B) \neq \varnothing$, there exists $y \in C(B)$.
- So $x \succsim^* y$. Hence, by WARP, $x \in C(B)$. This proves

$$C^*(B, \succsim^*) \subseteq C(B) \text{ for all } B \in \mathfrak{B}.$$

- These two set inclusions imply the desired conclusion: $C(B) = C^*(B, \succsim^*)$ for all $B \in \mathfrak{B}$. ■

Proof of Part 2 of the Theorem (con't con't)

We now show that if $\succsim$ is reflexive and rationalizes $\langle \mathfrak{B}, C \rangle$, then $\succsim^* \subseteq \succsim$:

- Suppose $x \succsim^* y$ for some $x, y \in X$.
- If $x = y$ we are done: as $\succsim$ is reflexive, $x \succsim y$. So assume $x \neq y$.
- Then $x \succsim^* y$ implies $B \in \mathfrak{B}$ exists such that $y \in B$ and $x \in C(B)$.
- Since $C^*(B, \succsim) = C(B)$, we now have $x \in C^*(B, \succsim)$.
- Thus, as $y \in B$, $x \succsim y$. $\blacksquare$

Lastly, we show $\succsim \subseteq \succsim^*$ if $\succsim$ rationalizes $\langle \mathfrak{B}, C \rangle$ and $\mathfrak{B}$ contains all 2-sets.

- Suppose $x \succsim y$ for some $x, y \in X$.
- If $x = y$ we are done: as $\succsim^*$ is reflexive, $x \succsim^* y$. So assume $x \neq y$.
- Since $x \succsim y$, $x \in C^*(\{x, y\}, \succsim)$.
- By assumption, $\{x, y\} \in \mathfrak{B}$. So $C(B)$ is well defined.
- Since $\succsim$ rationalizes $\langle \mathfrak{B}, C \rangle$, $C(B) = C^*(\{x, y\}, \succsim)$.
- Therefore $x \in C(B)$, and so $x \succsim^* y$. $\blacksquare$

The As-If Hypothesis

Under the conditions of part 2 of the theorem, the DM's choice behavior is rationalized by the complete and transitive revealed preference relation $\succsim^*$.

Her behavior is **as if** $\succsim^*$ is her "true" psychological preference relation. Her **choice process** could be to find and choose in each feasible set an alternative that is maximal according to $\succsim^*$.

However, her actual choice process may be entirely different! And $\succsim^*$ may bear no resemblence to her actual psychological preferences!

Example (Herb Simon) Satisficing

$\preceq$ is a rational preference relation on $X = \{x_1, \ldots, x_n\}$

$x_s \in X$ is a designated "satisfactory" alternative

For each nonempty $B \subseteq X$, $C(B)$ is the alternative in B with the largest subscript if $x_s \succ x_i$ for all $x_i \in B$, and otherwise it is the alternative in B with the smallest subscript satisfying $x_i \succsim x_s$.

Claim. $\langle 2^X \backslash \varnothing, C \rangle$ satisfies WARP, and hence is rationalized by its rational revealed preference order, $\succsim^*$. However, in general,

$$\succsim^* \;\neq\; \succsim$$

is possible, in which case policy recommendations based on $\succsim^*$ might be very bad!

Example. Let the true prefs be $x_n \succ x_{n-1} \succ \cdots \succ x_1$, and the satisfactory alternative be x_1. Then $x_1 \succ^* x_2 \succ^* \cdots \succ^* x_n$.