

Decision Theory

Decision theory is about making choices

- It has a *normative* aspect
 - what “rational” people *should* do
- ...and a *descriptive* aspect
 - what people *do* do

Not surprisingly, it’s been studied by economists, psychologist, and philosophers.

More recently, computer scientists have looked at it too:

- How should we design robots that make reasonable decisions
- What about software agents acting on our behalf
 - agents bidding for you on eBay
 - managed health care
- Algorithmic issues in decision making

This course will focus on normative aspects, informed by a computer science perspective.

Axiomatic Decision Theory

Standard (mathematical) approach to decision theory:

- Give axioms characterizing reasonable decisions
 - Ones that any “rational” person should accept
- Then show that these axioms characterize a particular approach to decision making
 - For example, we will discuss Savage’s axioms that characterize maximizing expected utility.
- An issue that arises frequently:
 - How to *represent* a decision problem

Choice Under Certainty

Assumption: you're given a set X of objects.

- You have to state which one you want most
- More generally: you give a preference order among objects in X
 - Notation: $x \succ y$ means x is strictly preferred to y
- There's no uncertainty: you get what you choose.
- **Goal:** to understand reasonable properties of a preference order

Example: $X = \{a, b, c\}$, $b \succ a$, $a \succ c$, and $c \succ a$.

- Are such cyclic preferences reasonable?
- Would a rational person have such preferences?

Axioms for Choice Under Certainty

Asymmetry: If $x \succ y$ then $y \not\succeq x$.

Negative Transitivity: If $x \not\succeq y$ and $y \not\succeq z$ then $x \not\succeq z$

Transitivity: If $x \succ y$ and $y \succ z$, then $x \succ z$.

Proposition: Asymmetry + NT imply Transitivity.

Proof: Suppose that $x \succ y$, $y \succ z$, and $x \not\succeq z$. By asymmetry, we have $z \not\succeq y$. By NT, we have $x \not\succeq y$ —contradiction.

Is NT a good normative or descriptive property?

Proposition. The binary relation \succ is negatively transitive iff $x \succ z$ implies that, for all $y \in X$, $x \succ y$ or $y \succ z$. (Proof in Kreps)

- NT comes pretty close to saying that \succ is a total order
- For all x, y , either (1) $x \succ y$, (2) $y \succ x$, or (3) for all z , (a) $x \succ z$ iff $y \succ z$, and (b) $z \succ x$ iff $z \succ y$
 - in case (3), x and y are “equivalent”

Definition. A binary relation \succ is a (strict) *preference relation* if it is asymmetric and negatively transitive.

Weak Preference

Definition. For $x, y \in X$:

1. x is *weakly preferred* to y , $x \succeq y$, if $y \not\succeq x$.
2. x is *indifferent* to y , $x \sim y$, if $x \not\succeq y$ and $y \not\succeq x$.

Definition. The binary relation \succeq on X is *complete* if for all $x, y \in X$, $x \succeq y$, $y \succeq x$ or both. It is *transitive* if $x \succeq y$ and $y \succeq z$ implies $x \succeq z$.

Proposition. Let \succ be a binary relation on X .

1. \succ is asymmetric iff \succeq is complete.
2. \succ is negatively transitive iff \succeq is transitive.

Proof of \Rightarrow

1. Given $x, y \in X$, asymmetry implies that can't have both $x \succ y$ and $y \succ x$. So at least one of $x \not\succeq y$ and $y \not\succeq x$ holds. Thus either $y \succeq x$ or $x \succeq y$. This is completeness.
2. If $x \succeq y$ and $y \succeq z$, then $y \not\succeq x$ and $z \not\succeq y$. By NT, $z \not\succeq x$; i.e., $x \succeq z$. This is transitivity.

\Leftarrow will be on homework 1.

Transitivity

Why do we care about transitivity?

- Is it normative?
- People's preferences don't always satisfy transitivity, but if this is pointed out, most people think they should change their preferences (or think the problem is not adequately represented).

Transitivity seems important for choice.

Example. $X = \{a, b, c\}$. Consider a sequence of choices from among pairs.

1. $\{a, b\}$, $a \succ b$, so a is chosen.
2. $\{a, c\}$, $c \succ a$, so c is chosen.
3. $\{b, c\}$, $b \succ c$, so b is chosen.
4. $\{a, b\}$...

Without transitivity can get cycles.

Lemma: If \succ is a preference relation then \succ is acyclic, i.e. $[x_1 \succ x_2 \succ \dots \succ x_{n-1} \succ x_n] \Rightarrow [x_1 \neq x_n]$.

Choice

Extend binary comparisons to choice over a set of more objects.

- Let $\mathcal{P}^+(X)$ be the set of all non-empty subsets of X .

Definition. For \succ a preference relation on X define

$$c(A, \succ) = \{x \in A : \text{for all } y \in A, y \not\succeq x\}.$$

Interpretation: $c(A, \succ)$ is the set of alternatives chosen from A by a decision maker with preferences \succ .

- If $x, y \in c(A, \succ)$ then $x \sim y$.

Proposition. If \succ a preference relation on X , and A is a finite subset of X , then $\emptyset \neq c(A, \succ) \subseteq A$.

- $c(A, \succ)$ is a nonempty subset of A
- This depends on A being finite and \succ being acyclic

Properties of choice functions

Definition. A *choice function* for X is a function $c : \mathcal{P}^+ \rightarrow \mathcal{P}^+$ such that for all finite $A \in \mathcal{P}^+$, $c(A)$ is a nonempty subset of A .

- $c(\cdot, \succ)$ is a choice function.

Not every choice function is generated by some \succ .

Example. $X = \{a, b, c\}$.

1. $c(\{a, b, c\}) = \{a\}$ and $c(\{a, b\}) = \{b\} \Rightarrow$ a violation of asymmetry.
2. $c(\{a, b\}) = \{a, b\}$ and $c(\{a, b, c\}) = \{b\} \Rightarrow$ a violation of NT.

Axioms of Choice Functions

Sen's α . If $x \in B \subseteq A$ and $x \in c(A)$, then $x \in c(B)$.
(This is apparently due to Chernoff, not Sen.)

Proposition. If \succ is a preference relation then $c(\cdot, \succ)$ satisfies Sen's α .

Proof. Suppose there are sets $A, B \in \mathcal{P}^+$ with $x \in B \subseteq A$, $x \in c(A, \succ)$, and $x \notin c(B, \succ)$. Then there is a $y \in B$ such that $y \succ x$. Since $B \subseteq A$ we have $y \in A$ and $y \succ x$. Thus $x \notin c(A, \succ)$ —contradiction.

Sen's β . If $x, y \in c(A)$, $A \subseteq B$, and $y \in c(B)$, then $x \in c(B)$.

Proposition. If \succ is a preference relation then $c(\cdot, \succ)$ satisfies Sen's β .

Proof. Since $x \in c(A, \succ)$ and $y \in A$ we have $y \not\succeq x$. By definition, $y \in c(B, \succ)$ implies that for all $z \in B$, $z \not\succeq y$. By negative transitivity, $y \not\succeq x$ and $z \not\succeq y$ implies $z \not\succeq x$. Since $x \in B$ and this holds for all $z \in B$ we have $x \in c(B, \succ)$.

Are there other restrictions on $c(\cdot, \succ)$ that follow from \succ being a preference relation?

- No!

Proposition. If a choice function c satisfies Sen's α and β , then there is a preference relation \succ such that $c(\cdot) = c(\cdot, \succ)$.

Proof: Define the “revealed preference” relation \succ by

$$x \succ y \text{ if } x \neq y \text{ and } c(\{x, y\}) = \{x\}.$$

To prove the proposition we need to show that \succ is a preference relation and that $c(\cdot) = c(\cdot, \succ)$.

\succ is a preference relation:

1. Asymmetry. Suppose $x \succ y$ and $y \succ x$. Then $c(\{x, y\}) = \{x\}$ and $c(\{x, y\}) = \{y\}$ —contradiction.
2. Negative Transitivity. Suppose that $z \not\succeq y$ and $y \not\succeq x$. We need to show that $z \not\succeq x$; i.e., $x \in c(\{x, z\})$. By Sen's α , showing that $x \in c(\{x, y, z\})$ is sufficient. Suppose $x \notin c(\{x, y, z\})$. Then at least one of y and z are in $c(\{x, y, z\})$.
 - If $y \in c(\{x, y, z\})$, then by Sen's α , $y \in c(\{x, y\})$. Since $y \not\succeq x$, cannot have $c(\{x, y\}) = \{y\}$. Thus, $x \in c(\{x, y\})$. By Sen's β , $x \in c(\{x, y, z\})$ —contradiction.

- If $z \in c(\{x, y, z\})$, then by Sen's α , $z \in c(\{y, z\})$. Since $z \neq y$, we must have $y \in c(\{y, z\})$. By Sen's β , $y \in c(\{x, y, z\})$ —contradiction, by the previous argument.

Now we show that $c(A) = c(A, \succ)$ for all $A \in \mathcal{P}^+$,

1. Suppose $x \in c(A)$. Then by Sen's α , $x \in c(\{x, y\})$ for all $y \in A$. Thus for all $y \in A$, $y \neq x$. So $x \in c(A, \succ)$.
2. Suppose $x \in c(A, \succ)$. Then for all $y \in A$, $y \neq x$. So for all $y \in A$, $x \in c(\{x, y\})$. Suppose $x \notin c(A)$. Then there is some $z \in A$, $z \neq x$ such that $z \in c(A)$. By Sen's α , $z \in c(\{x, z\})$. Then $c(\{x, z\}) = \{x, z\}$, $\{x, z\} \subseteq A$ and $z \in c(A)$. So by Sen's β , $x \in c(A)$ —contradiction.

So Sen's α and β completely characterize the choice functions that can be represented by a preference relation; i.e., those choice functions of the form $c(\cdot, \succ)$.

- This is an example of a representation theorem!
 - There are many others in the literature
 - we'll see a few of them

WARP

Houthaker's Axiom is an alternative to Sen's α and β .

- also called the Weak Axiom of Revealed Preference (WARP).

WARP: If $x, y \in A \cap B$, $x \in c(A)$, and $y \in c(B)$, then $x \in c(B)$ and $y \in c(A)$.

Theorem: $c(\cdot)$ satisfies Sen's α and β if and only if it satisfies WARP.

Proof: Assume $c(\cdot)$ satisfies WARP.

- Sen's α is a special case where $x = y$
- Sen's β is also clearly a special case

Assume $c(\cdot)$ satisfies Sen's α and β

- Then c is representable by some preference order \succ
- So c clearly satisfies WARP

Partial Orders

Completeness is questionable from both a descriptive and a normative point of view.

- Why should a decision maker always be able to compare x and y
- Technically, NT may fail

Definition. \succ is a *partial order* if it is an asymmetric and transitive binary relation.

We can define a choice function as before. What properties does it have?

- Sen's α still holds, but Sen's β may fail.
- (See homework 1.)

Now we would not want to define \sim as before.

- $x \not\succeq y$ and $y \not\succeq x$ could express indifference or non-comparability.
- Need to describe both strict preference and indifference: i.e. preferences described by the pair (\succ, \sim) .
- Another alternative: just use \preceq