Recovering utility

C. Chambers F. Echenique N. Lambert
Georgetown UC Berkeley USC

RUD Kyoto University, June 23 2023

This paper:

- An experimenter (Alice) and a subject (Bob).
- Alice conducts a finite choice experiment of size k:

$$\{x_1, y_1\}, \ldots, \{x_k, y_k\};$$

  k questions, each one a binary choice problem.
- Bob makes choices according to some $\succeq^*$ with utility u^*.
- Preference $\succeq_k$, with utility u_k, obtained as rationalizations or estimates.

How are $\succeq_k$, $\succeq^*$, u_k and u^* related?

This paper:

We present some answers in two kinds of models.

A deterministic model, and a statistical model that allows for randomness in sampling and choice.

Motivation

Decision theory:

"Utility representation iff axioms. Moreover, utility is unique."

Axioms $\Rightarrow$ testable implications.

Uniqueness $\Rightarrow$ identification, which enables estimation, or recovery of, a utility function.

In both cases assuming infinite data.

Motivation

We:

- Want to understand the problem with finite data.
- But large sample size – big data :)
- Focus on recovery.

Our results provide sufficient conditions that enable utility recovery in large, but finite, samples. Including approximation guarantees in statistical models with "mistakes" and sampling noise.

Motivation

Theorem

Suppose that $\succeq$ and $\succeq^*$ are two continuous preference relations (complete and transitive w.o.) on $\mathbb{R}^d$ and $B \subseteq \mathbb{R}^d$ dense.
If $\succeq|_{B\times B} = \succeq^*|_{B\times B}$, then $\succeq = \succeq^*$.

With infinite data, there's no problem (for preferences). But it's easy to exhibit examples of incorrect inference with (arbitrarily large) finite data.

Motivation

Choice space $X = \mathbb{R}^d$.

Theorem (Informal)

Let

- $\succeq^*$ be monotone and cont.;
- $\succeq_k$ strongly rationalize the k-sized choice data generated by $\succeq^*$.

Then,

- $\succeq_k \to \succeq^*$ (in the topology of closed convergence).
- For any utility u^* for $\succeq^*$ $\exists$ u_k for $\succeq_k$ s.t $u_k \to u^*$ (in the topology of compact convergence).

Utilities are more delicate than preferences.

Literature

Afriat's theorem and revealed preference tests: Afriat (1967); Diewert (1973); Varian (1982); Matzkin (1991); Chavas and Cox (1993); Brown and Matzkin (1996); Forges and Minelli (2009); Carvajal, Deb, Fenske, and Quah (2013); Reny (2015); Nishimura, Ok, and Quah (2017)

Recoverability: Varian (1982); Cherchye, De Rock, and Vermeulen (2011), Chambers-Echenique-Lambert (2021).

Approximation: Mas-Colell (1978); Forges and Minelli (2009); Kübler and Polemarchakis (2017); Polemarchakis, Selden, and Song (2017)

Identification: Matzkin (2006); Gorno (2019)

Econometric methods: Matzkin (2003); Blundell, Browning, and Crawford (2008); Blundell, Kristensen, and Matzkin (2010); Halevy, Persitz, and Zrill (2018)

Model

- A topological space X.
- Preference: A complete and continuous binary relation $\succeq$ over X
- $\mathcal{P}$ a set of preferences.

A pair $(X, \mathcal{P})$ is a preference environment.

Example: Expected utility preferences

- There are d prizes.
- X is the set of lotteries over the prizes, $\Delta^{d-1} \subset \mathbb{R}^d$.
- An EU preference $\succeq$ is defined by $v \in \mathbb{R}^d$ such that $p \succeq p'$ iff $v \cdot p \geq v \cdot p'$.
- $\mathcal{P}$ is set of all the EU preferences.

Local strictness

A preference $\succeq$ is *locally strict* if, for all $x, y \in X$, $x \succeq y$ implies that for each nbd U of (x, y), there is $(x', y') \in U$ with $x \succ y$.

Introduced by Border and Segal (1994) as a generalization of local non-satiation.

Experiment

Alice wants to recover Bob's preference from his choices.

- Binary choice problem: $\{x, y\} \subset X$.
- Bob is asked to choose x or y.
  Behavior encoded in a choice function $c(\{x, y\}) \in \{x, y\}$.
- If Bob's preference is $\succeq$ then $c(\{x, y\}) \succeq x$ and $c(\{x, y\}) \succeq y$.

Experiment

Alice gets finite dataset.

- Experiment of size k : $\Sigma^k = \{\{x_1, y_1\}, \ldots, \{x_k, y_k\}\}$.
- Set of growing experiments: $\{\Sigma^k\}$with $\Sigma^k \subset \Sigma^{k+1}$.

To sum up:

- $(X, \mathcal{P})$ preference env.
- c encodes choice
- Σ^k seq. of experiments

Topology on preferences

Choice of topology: closed convergence topology.

- Standard topology on preferences (Kannai, 1970; Mertens (1970); Hildenbrand, 1970).
- $\succeq_n \to \succeq$ when:
  1. For all $(x, y) \in \succeq$, there exists a seq. $(x_n, y_n) \in \succ_n$ that converges to (x, y).
  2. If a subsequence $(x_{n_k}, y_{n_k}) \in \succeq_{n_k}$ converges, the limit belongs to $\succeq$.
- If X is compact and metrizable, same as convergence under the Hausdorff metric.
- X Euclidean and $\mathcal{B}$ the strict parts of cont. weak orders. Then it's the smallest topology for which the set

$$\{(x, y, \succ) : x \in X, y \in X, \succ \in \mathcal{B} \text{ and } x \succ y\}$$

is open.

Topology on preferences

Lemma

Let X be a locally-compact Polish space. Then the set of all continuous binary relations on X is a compact metrizable space.

Topology of compact convergence

Let $\{u_k\}$ be a sequence of functions,

$$u_k \colon X \to \mathbb{R}.$$

The sequence *convergences compactly* to $u \colon X \to \mathbb{R}$ if for every compact $K \subseteq X$,

$$u_k|_K \to u|_K$$

uniformly.

Turn out to be the right topology for utility functions when preferences are endowed with the closed convergence topology.

Recovery of utility functions

Standard representation

Finite state space: S.

Monetary consequences: $[a, b] \subseteq \mathbb{R}$

Anscombe-Aumann acts: $f : S \to \Delta([a, b])$

Preferences on $\Delta([a, b])^S$.

Standard representation

Let U be the set of all continuous and monotone weakly increasing functions $u : [a, b] \to \mathbb{R}$ with $u(a) = 0$ and $u(b) = 1$.

A pair (V, u) is a *standard representation* if $V : \Delta([a, b])^S \to \mathbb{R}$ and $u \in U$ are continuous functions such that $v(p, \ldots, p) = \int_{[a,b]} u \,\mathrm{d}p$, for all constant acts $(p, \ldots, p)$.

(V, u) is *aggregative* if there is an *aggregator* $H : [0, 1]^S \to \mathbb{R}$ with $V(f) = H((\int u \,\mathrm{d}f(s))_{s \in S})$ for $f \in \Delta([a, b])^S$.

An aggregative representation with aggregator H is denoted by (V, u, H).

Standard representation

A preference $\succeq$ on $\Delta([a, b])^S$ is *standard* if it is weakly monotone, and there is a standard representation (V, u) in which V represents $\succeq$.

Example

Variational preferences (Maccheroni et al 2006) are standard and aggregative. Let

$$V(f) = \inf\{\int v(f(s))d\pi(s) + c(\pi) : \pi \in \Delta(S)\}$$

where

1. $v : \Delta([a, b]) \to \mathbb{R}$ is continuous and affine.
2. $c : \Delta(S) \to [0, \infty]$ is lower semicontinuous, convex and grounded (meaning that $\inf\{c(\pi) : \pi \in \Delta(S)\} = 0$).

Let $H : [0, 1]^S \to \mathbb{R}$ be $H(x) = \inf\{\sum_{s\in S} x(s)\pi(s) + c(\pi) : \pi \in \Delta(S)\}$

Standard representation

Theorem

Let $\succeq$ be a standard preference with standard representation (V, u), and $\{\succeq^k\}$ a sequence of standard preferences, each with a standard representation (V^k, u^k).

1. If $\succeq^k \to \succeq$, then $(V^k, u^k) \to (V, u)$.
2. If, in addition, these preferences are aggregative with representations (V^k, u^k, H^k) and (V, u, H), then $H^k \to H$.

Statistical model

Choice among vectors in $\mathbb{R}^d$.
Focus on the *Wald representation* of $\succeq$, $u : \mathbb{R}^d \to \mathbb{R}$ s.t

$$x \sim (u(x), \ldots, u(x)).$$

Statistical model

Primitives $(X, \mathcal{P}, \lambda, q)$:

- $X \subseteq \mathbb{R}^d$ is the choice space.
- $\mathcal{P}$ is a class of cont. and l.s. preferences on X. Comes with a set of Wald utility functions $\mathcal{U}$, so each preference in $\mathcal{P}$ has a Wald representation in $\mathcal{U}$.
- λ is a (Borel) probability measure on X.
- $q : X \times X \times \mathcal{P} \to [0, 1]$ is a random choice function, so $q(x, y; \succeq)$ is the probability that an agent with preferences $\succeq$ chooses x over y.

Assumptions

- λ is abs. cont. with respect to Lebesgue measure, and satisfies $\lambda \geq c\,\mathrm{Leb}$, where $c > 0$ is a constant.
- If $x \succ y$, then x is chosen with probability $q(x, y; \succeq) > 1/2$ and y with probability $q(y, x; \succeq^*) = 1 - q(x, y; \succeq)$. If $x \sim y$ then x and y are chosen with equal probability.
- $$\Theta \equiv \inf\{q(\succeq, (x, y)) : x \succ y \text{ and } \succeq \in \mathcal{P}\} > \frac{1}{2}.$$
- The space of utility functions is endowed with a metric, ρ.

Estimator

Given a dataset of size k.

u_k is chosen to max the number of rationalized choices in the data.

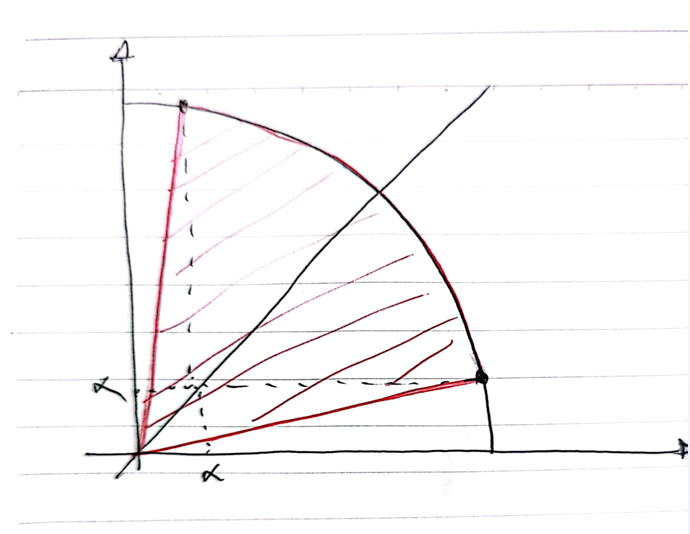
Lipschitz

$(X, \mathcal{P}, \lambda, q)$ is a *Lipschitz environment* if:

1. $X \subseteq \mathbb{R}^d$ is convex, compact, and has nonempty interior.
2. All utilities in $\mathcal{U}$ are Lipschitz w/common Lip. constant κ.

Homothetic

Notation: $S_\alpha^M = \{x \in \mathbb{R}^d : \|x\| = M \text{ and } x \geq \alpha 1\}$ and $D_\alpha^M = \{\theta x : x \in S_\alpha^M \text{ and } \theta \in [0, 1]\}$.



Homothetic

$(X, \mathcal{P}, \lambda, q)$ is a *homothetic environment* if:

1. $X = D_\alpha^M$ for some (small) $\alpha > 0$ and (large) $M > 0$.
2. $\mathcal{P}$ is a class of cont., monotone, homothetic, and complete preferences on $X \subseteq \mathbb{R}^d$.

VC dimension

The VC dimension of $\mathcal{P}$ is the largest cardinality of an experiment that can always be rationalized by $\mathcal{P}$.

A measure of how flexible $\mathcal{P}$; how prone it is to overfitting.

VC dimension

- Think of a game between Alicia and Roberto
- Alicia defends $\mathcal{P}$; Roberto questions it.
- Given is k
- Alicia proposes a choice experiment of size k
- Roberto fills in choices adversarily.
- Alicia wins if she can rationalize the choices using $\mathcal{P}$.
- The VC dimension of $\mathcal{P}$ is the largest k for which Alicia always wins.

Convergence

Theorem

Let $(X, \mathcal{P}, \lambda, q)$ be either homothetic or Lipschitz. Suppose that $u^* \in \mathcal{U}$ is the Wald utility representation of $\succeq^* \in \mathcal{P}$.

1. Then estimates u_k converge to u^* in probability.
2. $\exists$ constants K and $\bar{C}$ s.t, for any $\delta \in (0,1)$ and k, w/prob. $\geq 1 - \delta$:

$$\rho(u_k, u^*) \leq \bar{C}\left(K\sqrt{V/n} + \sqrt{2\ln(1/\delta)/n}\right)^{1/D},$$

where V is the VC dimension of $\mathcal{P}$, $D = d$ when the environment is Lipschitz and $D = 2d$ when it is homothetic.

Ideas

$\mu(\succeq'|\succeq)$ is the prob that a random binary comparison from preference $\succeq$ (and λ and q) is consistent with $\succeq'$.

Key identification lemma: $\succeq' \neq \succeq$ implies $\mu(\succeq'|\succeq) < \mu(\succeq|\succeq)$.

As a consequence, if u_k is maximizing an objective that is the sample analogue of μ, when the sample is large (we show) the preference it represents can't be too far from the one generating the choices.

Ideas

The assumptions on $(X, \mathcal{P}, \lambda, q)$ serve to connect $\rho(u, u')$ with $\mu(\succeq'|\succeq)$.

For ex.

Lemma

Consider a Lipschitz noise choice environment $(X, \mathcal{P}, \lambda, q)$. There is a constant C with the following property. If $\succeq$ and $\succeq'$ are two preferences in $\mathcal{P}$ with representations u and u' (respectively) in $\mathcal{U}$. Then

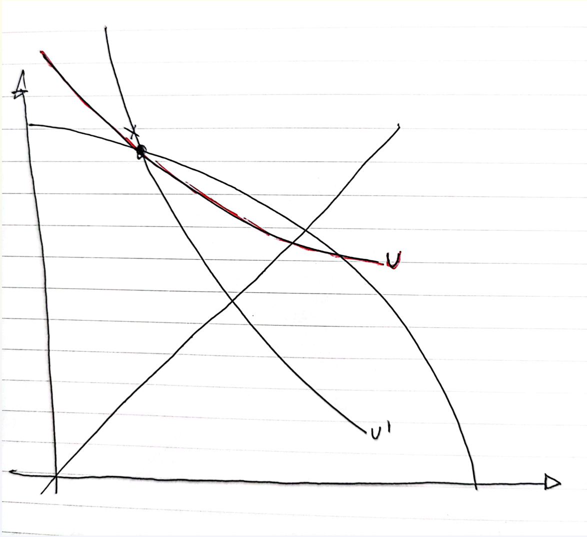
$$C\rho(u, u')^d \leq \mu(\succeq, \succeq) - \mu(\succeq', \succeq)$$

Lemma

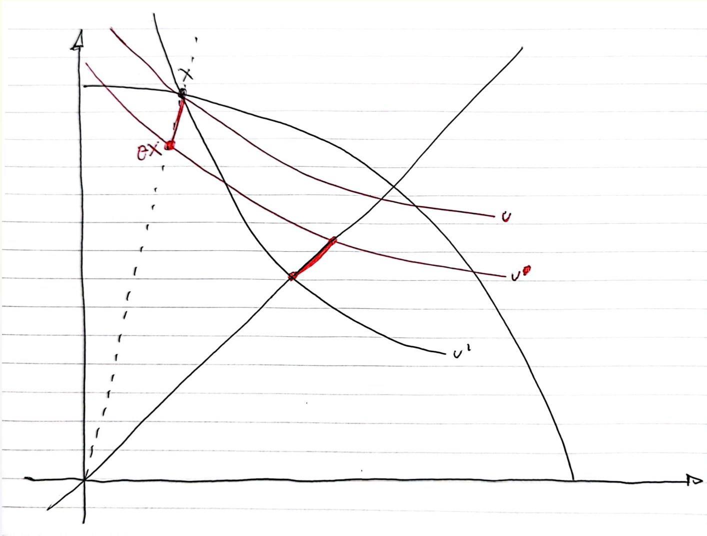
Consider a homothetic noise choice environment $(X, \mathcal{P}, \lambda, q)$. There is a constant C with the following property. If $\succeq$ and $\succeq'$ are two preferences in $\mathcal{P}$ with representations u and u' (respectively) in $\mathcal{U}$. Then

$$C\rho(u, u')^{2d} \leq \mu(\succeq, \succeq) - \mu(\succeq', \succeq)$$

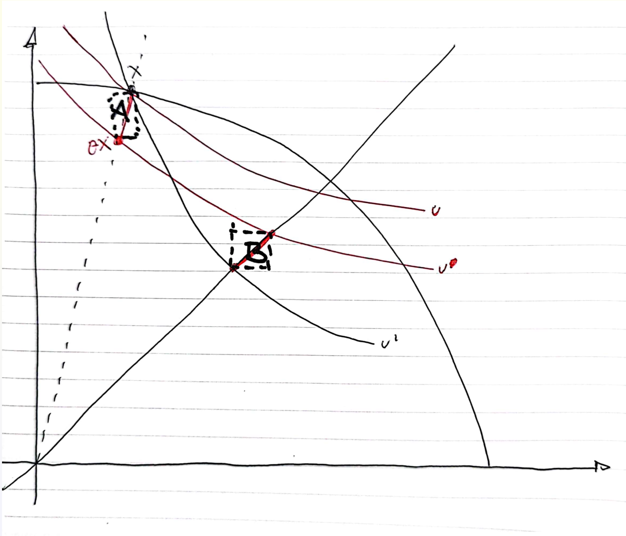
Ideas



Ideas



Ideas



Conclusion

- Binary choice
- Finite data
- "Consistency" – Large sample theory
- Unified framework: RP and econometrics.

Applicable to:

Large-scale (online) experiments/surveys.

Voting (roll-call data).