Recovering Preferences from Finite Data

Christopher Chambers[1], Federico Echenique[2], Nicolas Lambert[3]

[1]Georgetown University

[2]California Institute of Technology

[3]MIT

NYU Theory Workshop

October 7th 2020

This paper

- In a *revealed preference* model: When can we uniquely recover the data-generating preference as the dataset grows large?
- In an *statistical model*: Propose a consistent estimator.
- Unifying framework for both.

Applications:

- Expected utility preferences.
- Intertemporal consumption with discounted utility.
- Choice on commodity bundles.
- Choice over menus.
- Choice over dated rewards.
- ...

Model

Alice (an experimenter)

Bob (a subject)

Model

- Alice presents Bob with choice problems:

"Hey Bob would you like x or y?"

x vs. y

- Bob chooses one alternative.
- Rinse and repeat $\rightarrow$ dataset of n choices.

Model

- Alternatives: 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.

Examples

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.

Preferences on commodity bundles:

- There are d commodities.
- $X \equiv \mathbb{R}^d_+$, the i-th entry of a vector is quantity consumed of i-th good.
- $\mathcal{P}$ is set of all monotone preferences on X.

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 by a choice function $c(\{x, y\}) \in \{x, y\}$.
- Partial observability: indifference is not observable.

Experiment

Alice gets finite dataset.

- Experiment of length n : $\Sigma_n = \{B_1, \ldots, B_n\}$ with $B_k = \{x_k, y_k\}$.
- Set of growing experiments: $\{\Sigma_n\} = \{\Sigma_1, \Sigma_2, \ldots\}$ with $\Sigma_n \subset \Sigma_{n+1}$.

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)

Consistency: 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)

What's new?

Unified framework: rev. pref. and econometrics.

What's new?

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

OK, so far:

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

Rationalization/ Estimation

- Revealed Preference: A preference $\succeq$ rationalizes the observed choices on Σ_n if $\{x, y\} \in \Sigma_n$, $c(\{x, y\}) \succeq x$ and $c(\{x, y\}) \succeq y$.
- Statistical model: preference estimate ...

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

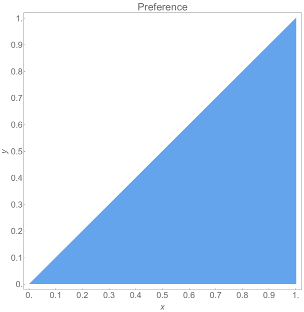
$$\{(x, y, \succ) : x \in X, y \in X, \succ \in \mathcal{B} \text{ and } x \succ y\}$$

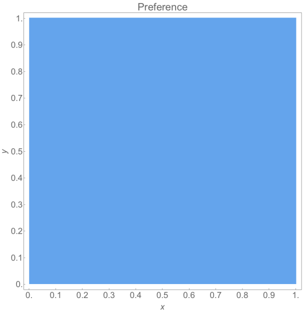
is open.

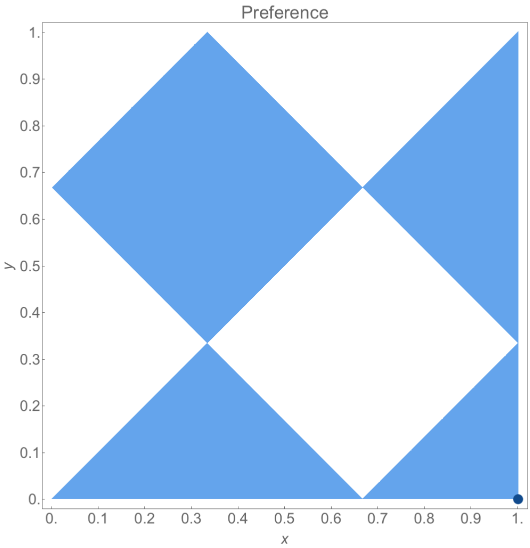
Examples

Set of alternatives $X = [0, 1]$.

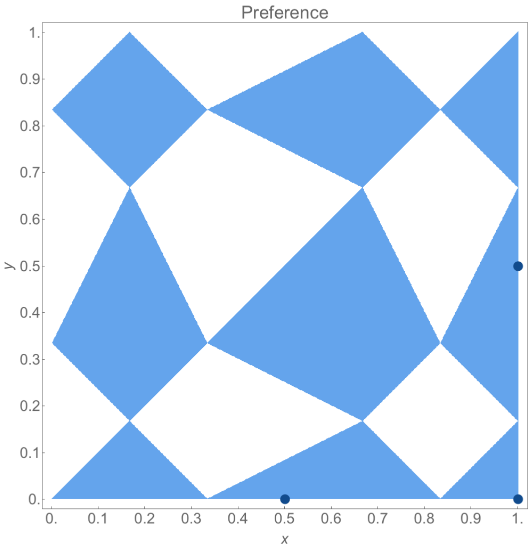
- Left: the subject prefers x to y iff $x \geq y$.
- Right: the subject is completely indifferent.



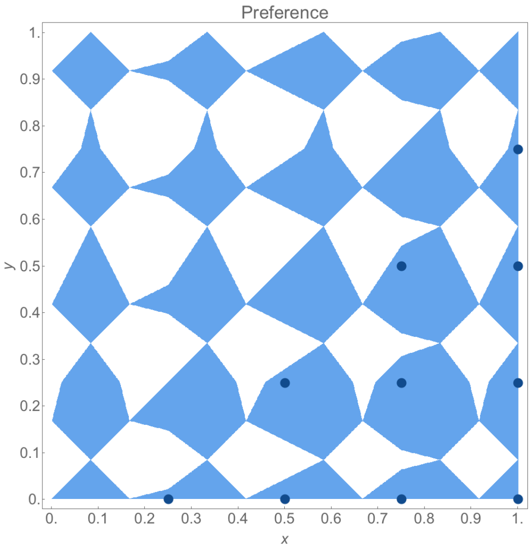



Preference
1.
0.9
0.8
0.7
0.6
y
0.5
0.4
0.3
0.2
0.1
0.
0.
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.
x

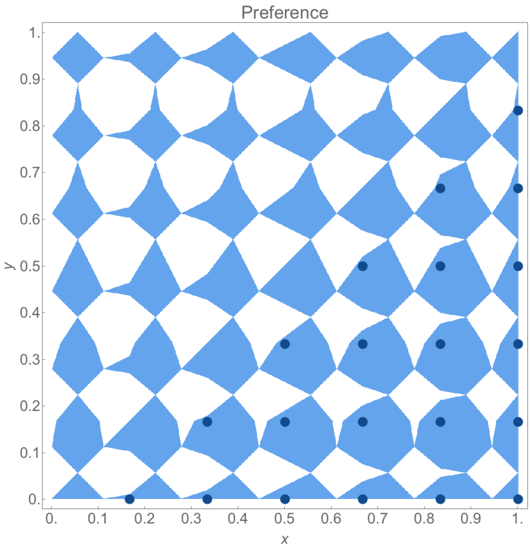
n=1



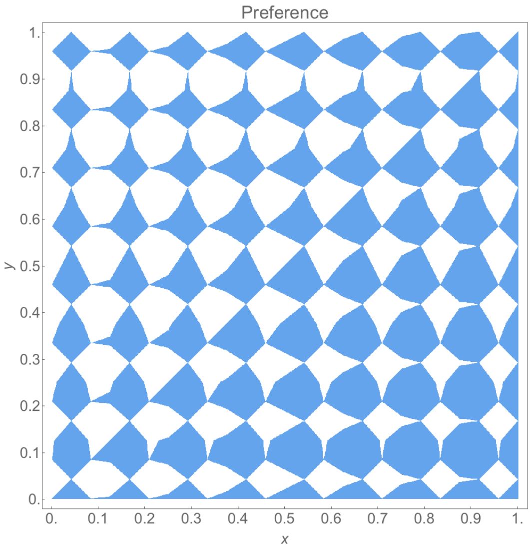
n=2


Preference
1.
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.
y
0.
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.
x

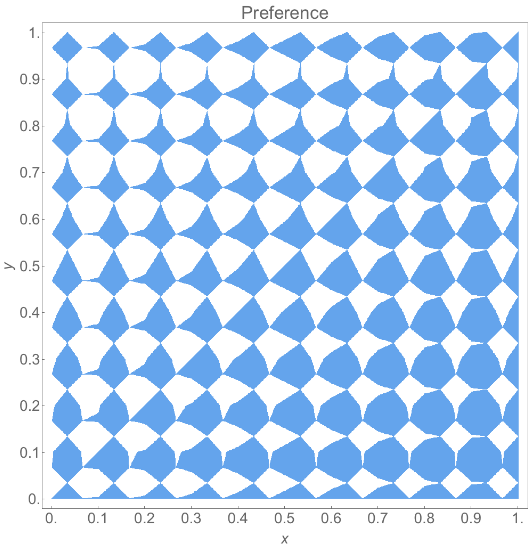
n=4


Preference
1.
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.
y
0.
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.
x

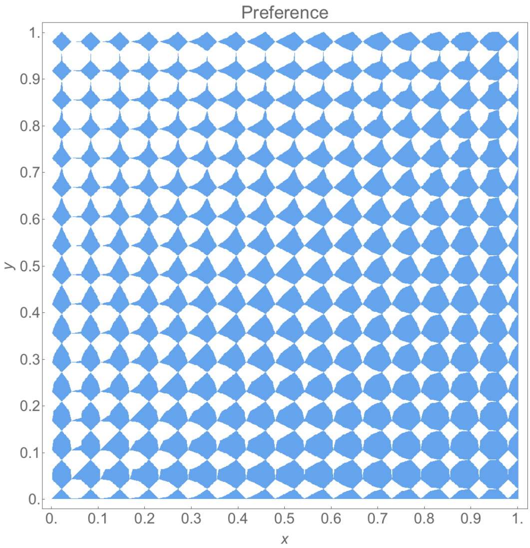
n=6


Preference
1.
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.
y
0.
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.
x

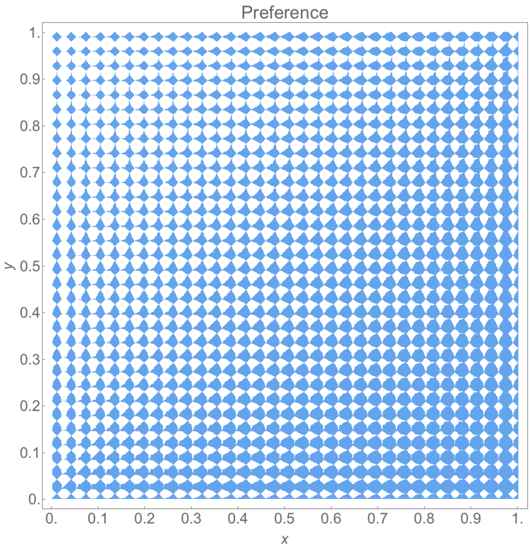
n=8


Preference
1.
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.
y
0.
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.
x

n=10


Preference
1.
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.
y
0.
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.
x

n=16

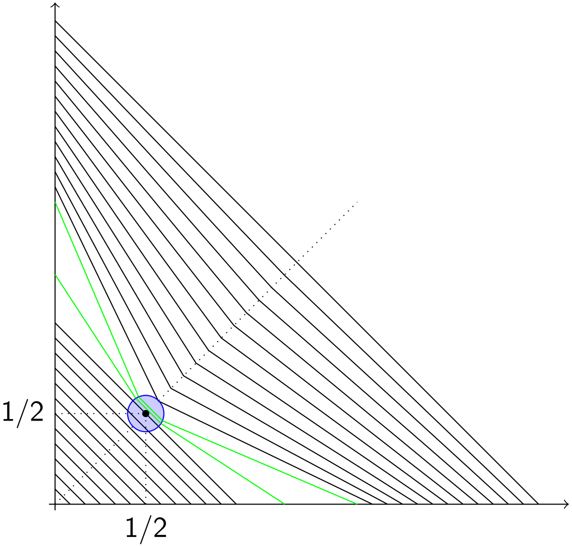

Preference
y
x
1.
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.
0.
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.

n=32

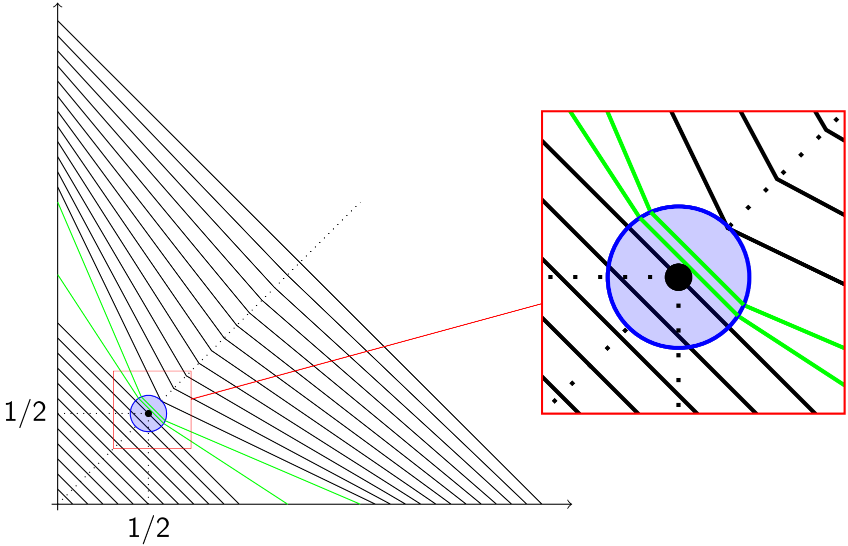
Moral

Discipline matters.

Non-closed $\mathcal{P}$



Non-closed $\mathcal{P}$



Moral

$\mathcal{P}$ must be closed, and some standard models are *not* closed.

Assumption on the set of alternatives

Assumption 1 : X is a locally compact, separable, and completely metrizable space.

Topology on preferences

Lemma

The set of all continuous binary relations on X is a compact metrizable space.

Assumption on the class of preferences

$\succeq$ is locally strict if

$x \succeq y \Longrightarrow$ in every nbd. of (x, y), there exists (x', y') with $x' \succ y'$

(Border and Segal, 1994).

Assumption on the class of preferences

Assumption 2 : $\mathcal{P}$ is a closed set of locally strict preferences.

Assumption on the set of experiments

A set of experiments $\{\Sigma_n\}$, with $\Sigma_n = \{B_1, \ldots, B_n\}$, is exhaustive when:

1. $\bigcup_{k=1}^{\infty} B_k$ is dense in X.
2. For all $x, y \in \bigcup_{k=1}^{\infty} B_k$ with $x \neq y$, there exists k such that $B_k = \{x, y\}$.

Assumption 3 : $\{\Sigma_n\}$ is an exhaustive growing set of experiments.

To sum up:

Assumption 1 : X is a locally compact, separable, and completely metrizable space.

Assumption 2 : $\mathcal{P}$ is a closed set of locally strict preferences.

Assumption 3 : $\{\Sigma_n\}$ is an exhaustive growing set of experiments.

First main result

Theorem 1

Suppose c is an arbitrary choice function.
When Assumptions (1), (2) and (3) are satisfied:

1. If, for every n, the preference $\succeq_n \in \mathcal{P}$ rationalizes the observed choices on Σ_n, then there exists a preference $\succeq^* \in \mathcal{P}$ such that $\succeq_n \to \succeq^*$.
2. The limiting preference is unique: if, for every n, $\succeq'_n \in P$ rationalizes the observed choices on Σ_n, then the same limit $\succeq'_n \to \succeq^*$ obtains.

So, if the subject chooses according to some preference $\succeq^* \in \mathcal{P}$, then $\succeq_n \to \succeq^*$.

Ideas behind the thm

Lemma

The set of all continuous binary relations on X is a compact metrizable space.

Lemma

If $A \subseteq X \times X$, then $\{\succeq \in X \times X : A \subseteq \succeq\}$ is closed.

Identification

Lemma

Consider an exhaustive set of experiments with binary choice problems $\{x_k, y_k\}$, $k \in \mathbb{N}$. Let $\succeq$ be any complete binary relation, and $\succeq_A$ and $\succeq_B$ be locally strict preferences. If, for all k, $x_k \succeq_A y_k$ and $x_k \succeq_B y_k$ whenever $x_k \succeq y_k$, then $\succeq_A = \succeq_B$.

Statistical model

Given $(X, \mathcal{P})$. We change:

- How subjects make choices: they do not exactly follow a preference, but randomly deviate from it.
- How experiments are generated.

Statistical model

1. In a choice problem, alternatives drawn iid according to sampling distribution λ.

2. Subjects make "mistakes."
   Upon deciding on $\{x, y\}$, a subject with preference $\succeq$ chooses x over y with probability $q(\succeq; x, y)$ (error probability function).

3. Only assumption: if $x \succ y$ then $q(\succeq; x, y) > 1/2$.

4. "Spatial" dependence of q on x and y is arbitrary.

Estimator

Kemeny-minimizing estimator: find a preference in $\mathcal{P}$ that minimizes the number of observations inconsistent with the preference.

- "Model free:" to compute estimator don't need to assume a specific q or λ.
- May be computationally challenging (depending on $\mathcal{P}$).

Assumption on the sampling distribution λ

Assumption 3' : λ has full support and for all $\succeq \in \mathcal{P}$,
$\{(x, y) : x \sim y\}$ has λ-probability 0.

Second main result

Theorem 2 (Part A)

Under Assumptions (1), (2), (3'), if the subject's preference is $\succeq^* \in \mathcal{P}$ and $\succeq_n$ is the Kemeny-minimizing estimator for Σ_n, then, $\succeq_n \to \succeq^*$ in probability.

Finite data

- Our paper is about finite data.
- Finite data but large samples
- How large?

Convergence rates: Digression

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.

Convergence rates: Digression

- 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 rates

- Let ρ be a metric on preferences.

Theorem 2 (Part B)

Under the same conditions as in Part A,

$$N(\eta, \delta) \leq \frac{2}{r(\eta)^2} \left(\sqrt{2/\delta} + C\sqrt{\mathrm{VC}(\mathcal{P})} \right)^2$$

Convergence rates

- Let ρ be a metric on preferences.
- $N(\eta, \delta)$: smallest value of N such that for all $n \geq N$, and all subject preferences $\succeq^* \in \mathcal{P}$,

$$\mathbf{Pr}(\rho(\succeq_n, \succeq^*) < \eta) \geq 1 - \delta.$$

Theorem 2 (Part B)

Under the same conditions as in Part A,

$$N(\eta, \delta) \leq \frac{2}{r(\eta)^2} \left(\sqrt{2/\delta} + C\sqrt{\mathrm{VC}(\mathcal{P})} \right)^2$$

Convergence rates

- Let ρ be a metric on preferences.
- $N(\eta, \delta)$: smallest value of N such that for all $n \geq N$, and all subject preferences $\succeq^* \in \mathcal{P}$,

$$\mathbf{Pr}(\rho(\succeq_n, \succeq^*) < \eta) \geq 1 - \delta.$$

- $\mu(\succeq'; \succeq)$: probability that the choice of a subject with preference $\succeq$ is consistent with preference $\succeq'$.

$$r(\eta) = \inf \left\{ \mu(\succeq; \succeq) - \mu(\succeq'; \succeq) : \succeq, \succeq' \in \mathcal{P}, \rho(\succeq, \succeq') \geq \eta \right\}.$$

Theorem 2 (Part B)

Under the same conditions as in Part A,

$$N(\eta, \delta) \leq \frac{2}{r(\eta)^2} \left(\sqrt{2/\delta} + C\sqrt{\mathrm{VC}(\mathcal{P})} \right)^2$$

Convergence rates

- Let ρ be a metric on preferences.
- $N(\eta, \delta)$: smallest value of N such that for all $n \geq N$, and all subject preferences $\succeq^* \in \mathcal{P}$,

$$\mathbf{Pr}(\rho(\succeq_n, \succeq^*) < \eta) \geq 1 - \delta.$$

- $\mu(\succeq'; \succeq)$: probability that the choice of a subject with preference $\succeq$ is consistent with preference $\succeq'$.

$$r(\eta) = \inf \left\{ \mu(\succeq; \succeq) - \mu(\succeq'; \succeq) : \succeq, \succeq' \in \mathcal{P}, \rho(\succeq, \succeq') \geq \eta \right\}.$$

- $\mathrm{VC}(\mathcal{P})$ the VC dimension of the class $\mathcal{P}$.

Theorem 2 (Part B)

Under the same conditions as in Part A,

$$N(\eta, \delta) \leq \frac{2}{r(\eta)^2} \left(\sqrt{2/\delta} + C \sqrt{\mathrm{VC}(\mathcal{P})} \right)^2$$

Expected utility

1. X is the set of lotteries over d prizes.
2. $\mathcal{P}$ is the set of nonconstant EU preferences: there are always lotteries p, p' such as p is strictly preferred to p'.

This preference environment satisfies Assumptions 1 and 2.

Suppose: there is $C > 0$ and $k > 0$ s.t

$$q(x, y; \succeq) \geq \frac{1}{2} + C(v \cdot x - v \cdot y)^k,$$

when $x \succeq y$ and v represents $\succeq$.

Expected utility

Under these assumptions, we can bound $r(\eta)$ and $\mathrm{VC}(\mathcal{P})$, which implies

$$N(\eta, \delta) = O\left(\frac{1}{\delta \eta^{4d-2}}\right).$$

Other examples: Cobb-Douglas, CES, and CARA subjective EU preferences, and intertemporal choice with discounted, Lipschitz-bounded utilities.

Monotone preferences

- K be a compact set in $X \equiv \mathbb{R}^d_{++}$, and fix $\theta > 0$.
- $\mathcal{P}$ has finite VC-dimension and is identified on K
- λ is the uniform probability measure on $K^{\theta/2}$,
- q satisfies: probability of choosing y instead of x when $x \succ y$ is a function of $\|x - y\|$,

Proposition

The Kemeny-minimizing estimator is consistent and, as $\eta \to 0$ and $\delta \to 0$,

$$N(\eta, \delta) = O\left(\frac{1}{\eta^{2d+2}} \ln \frac{1}{\delta}\right).$$

Applications: preferences from utilities

A set $\mathcal{P}$ is defined fom utilities when there is a class $\mathcal{U}$ of utility functions such that for all $\succeq \in \mathcal{P}$

$$x \succeq y \qquad \Leftrightarrow \qquad U(x) \geq U(y)$$

for some $U \in \mathcal{U}$.

Proposition 1

Under Assumption 1, if $\mathcal{U}$ is compact and represents locally strict preferences, then Assumption 2 is met.

Implied by the continuity theorem of Border and Segal (1994).

Revisit the case of expected utility preferences:

1. X is the set of lotteries over d prizes.
2. $\mathcal{P}$ is the set of nonconstant EU preferences: there are always lotteries p, p' such as p is strictly preferred to p'.

This preference environment satisfies Assumptions 1 and 2. When the probability of error of choosing y instead of x when $x \succ y$ is a function of $\|x - y\|$, we can bound $r(\eta)$ and $\mathrm{VC}(\mathcal{P})$, which implies

$$N(\eta, \delta) = O\left(\frac{1}{\delta \eta^{4d-2}}\right).$$

Other examples: Cobb-Douglas, CES, and CARA subjective EU preferences, and intertemporal choice with discounted, Lipschitz-bounded utilities.

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)

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)

Applications: monotone preferences

- Call a dominance relation any binary relation on X that is not reflexive.
- Say that $\succeq$ is strictly monotone wrt $\rhd$ if $x \rhd y$ implies $x \succ y$.
- Say that $\succeq$ is Grodal-transitive if $x \succeq y \succ z \succeq w$ implies $x \succeq w$.

Proposition 2

Take a set of alternatives X that meets Assumption 1, and suppose:

1. $\rhd$ is a dominance relation that is open,
2. for each x, there are y, z arbitrarily close to x such that $y \rhd x$ and $x \rhd z$.

Then the class of preferences that are Grodal-transitive and strictly monotone wrt $\rhd$ meets Assumption 2.

Example: back to preferences over commodity bundles.

- There are d commodities.
- $X \equiv \mathbb{R}^d_{++}$, where for $(x_1, \ldots, x_d) \in X$, x_i is quantity of good i consumed.
- $x \gg y$ iff $x_i > y_i$ for all $i = 1, \ldots, d$.

The set of all preferences that are Grodal-transitive and strictly monotone wrt $\gg$ meets Assumption 2.

Other examples: choice over menus of lotteries, dated rewards, intertemporal consumption, non-EU choice over lotteries.

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).