Axiomatization of random utility models with missing data

Presentation at EC London 2023

Haruki Kono, Kota Saito, Alec Sandroni

November 15, 2024

Random Utility Model

- Consider a group of individuals making a choice among the finite set X of alternatives:
- We observe a percentage $\rho(D, x)$ of people who choose each alternative x from a subset $D \subseteq X$.
  - Example: $X = \{a, b, c, d\}$
  - $\rho(\{a, b, c, d\}, a) = 20\%$
  - $\rho(\{a, b, c\}, a) = 23\%$
  - $\vdots$
  - $\rho(\{c, d\}, d) = 3\%$.
- This type of data is very common in economics, for example market share data.

Random Utility Model

- One fundamental model to describe this kind of data is *Random utility model*.
- This model specifies a probability distribution over possible rankings over X
- Example:
  - 10% of individuals have ranking $a \succ b \succ c \succ d$
  - 0% of individuals have ranking $a \succ b \succ d \succ c$
  - $\vdots$
  - 20% of individuals have ranking $d \succ c \succ b \succ a$

Random Utility Model

- Falmagne (1978) provides a necessary and sufficient condition on the data ρ under which ρ is consistent with a random utility model, that is there exists a probability distribution μ over rankings $\succ$ on X:

$$\rho(D,x) = \mu\Big(\Big\{ \ \succ \ \Big| x \succ y \text{ for all } y \in D \setminus x\Big\}\Big)$$

- Falmagne (1978) assumes that data is complete. That is, we know choice frequency $\rho(D,x)$ of any alternative $x \in D$ from any subset D of X.
- In reality, however, sometimes the data is incomplete.

Example: Transportation

- Consider the following transportation methods:

$$X = \{\text{bus, train, walk, drive}\}.$$

- The government may be able to estimate the market share of public transportation methods (bus or train) based on the revenues.
- However, it may be difficult for the government to know whether a passenger drives or walks (unless the government conducts a survey).
- Choice frequencies of walk and drive may not be available.

Examples

- Market Shares of Private Companies
  - One definition of market share is the percentage of a company's total sales divided by the market's total sales.
  - However, private companies sometimes do not disclose their financial information including total sales.
  - Market share of private companies may not be observable.
- School Choice with Private Schools
  - Applicants submit their rankings among public schools but not private schools.
  - Choice frequencies of private schools may not be observables.

Our Question

- In this setup of missing data, what is a necessary and sufficient condition on the observable data ρ under which ρ is consistent with random utility model μ?
- It is known that obtaining a tight necessary and sufficient condition is very difficult for the case of incomplete data in general.
  - When choice frequency is observable only for binary sets, how to obtain a tight necessary and sufficient condition has been an open question since the 1980s despite continuous effort in math, psychology, and economics.
  - The tight characterization is know only for the case when the number of alternatives is less than eight.
- We found a tight necessary and sufficient condition for this setup of missing data.

Model

- X: a finite set of alternatives.
- X^*: a subset of X. (the set of unobservable alternatives)
  - Assume that the choice frequencies of elements of X^* are not observable (even if a choice set includes the alternatives).
  - Let $\tilde{X} = X \setminus X^*$ (the set of observable alternatives).
- $\mathcal{D} \subseteq 2^X$: a set of choice sets.
  - In this presentation for simplicity, assume $\mathcal{D} = 2^X \setminus \emptyset$.

Model

- $\tilde{\mathcal{M}} = \{(D,x) \in \mathcal{D} \times X | x \in D, x \in \tilde{X}, D \in \mathcal{D}\}$.
  - The choice frequency ρ over (D,x) is observable (i.e., defined) if and only if $(D,x) \in \tilde{\mathcal{M}}$.
- Note that this does not mean that we cannot know anything about choice frequencies of $x^* \in X^*$.
  - When $x^* \in X^*$ is the only one unobservable alternative in the choice set D , $\rho(D,x^*)$ can be calculated as

$$1 - \sum_{y \in D \setminus \{x^*\}} \rho(D,y).$$

Example (Transportation):

- The government may be able to estimate the market share of public transportation methods (bus or train) based on the revenues.
- However, it is sometimes difficult for the government to know fractions of people who drive or walk.
- In this case, $X = \{\text{bus, train, walk, drive}\}$ and $X^* = \{\text{walk, drive}\}$.
- We assume that depending on locations of homes, some options are not available:

$$\mathcal{D} = \Big\{\{w\}, \{w, b\}, \{w, t\}, \{w, b, t\}, \\ \{w, d\}, \{w, d, b\}, \{w, d, t\}, \{w, d, b, t\}\Big\}.$$

Random Utility Rationalization

Let $\mathcal{L}$ be the set of rankings on X.

Definition 1

An incomplete dataset ρ is random *utility (RU) rationalizable* if there exists $\mu \in \Delta(\mathcal{L})$ such that for any $(D, x) \in \tilde{\mathcal{M}}$,

$$\rho(D, x) = \mu(\succ \in \mathcal{L} \mid x \succ y \text{ for all } y \in D \setminus \{x\}).$$

Falmagne (1978) and Block-Marschak polynomial

Assuming $X^* = \emptyset$ and $\mathcal{D} = 2^X \setminus \emptyset$, Falmagne (1978) showed that the following statements are equivalent

(i) ρ is RU rationalizable

(ii) $K(\rho, D, x) \geq 0$ for all (D, x) such that $x \in D \in 2^X$, where

$$K(\rho, D, x) \equiv \sum_{E:E \supseteq D} (-1)^{|E \setminus D|} \rho(E, x).$$

- Example:
  - If $D = X \setminus \{y\}$, then $K(\rho, D, x) = \rho(D, x) - \rho(X, x) \geq 0$.
  - If $D = X \setminus \{y, z\}$, then $K(\rho, D, x) =$ $\rho(D, x) - \rho(D \cup y, x) - \rho(D \cup z, x) + \rho(X, x) \geq 0$.

Falmagne (1978) and Block-Marschak polynomial

- Even for an incomplete data, note that BM polynomial $K(\rho, D, x)$ can be calculated if $(D, x) \in \tilde{\mathcal{M}}$.
- Mobius inversion implies that if ρ is a random utility model (represented by μ), then

$$K(\rho, D, x) = \mu(\succ | z \succ x \succsim y \text{ for all } y \in D \text{ and all } z \notin D).$$

For simplicity, I write the right hand side as $\mu(\succ | D^c \succ x \succsim D)$.

Definition

A collection $\mathcal{C}$ of subsets of X is a *test collection* if there exist

- a nonempty set $A \subsetneq X \setminus X^*$ of observable alternatives and
- a collection $\mathcal{E} \subsetneq 2^{X^*}$ of sets of unobservables alternatives

such that

$$\mathcal{C} = \{A \cup E | E \in \mathcal{E}\}.$$

and $\mathcal{E}$ is an upper set:

$$D \in \mathcal{E}, D \subseteq E \Rightarrow E \in \mathcal{E}.$$

Theorem

(a) An incomplete dataset $\rho \in \Re_{+}^{\mathcal{M} \setminus \mathcal{M}^*}$ is RU-rationalizable if and only if the following two conditions hold:

- (i) for any $(D, x) \in \mathcal{M} \setminus \mathcal{M}^*$ such that $1 < |D| < |X|$, $K(\rho, D, x) \geq 0$,
- (ii) for any test collection $\mathcal{C} \subseteq \mathcal{D}$,

$$\left(\sum_{(D,x):D \in \mathcal{C}, D \cup x \notin \mathcal{C}} K(\rho, D \cup x, x) - \sum_{(F,y):F \notin \mathcal{C}, F \cup y \in \mathcal{C}} K(\rho, F \cup y, y) \right) \geq 0$$

Comments:

- BM polynomial $K(\rho, D, x)$ is computable based on observable data if and only if $(D, x) \notin \mathcal{M}^*$.
- Condition (ii) can also be tested based on observables.
- Condition (ii) has a very intuitive explanation based on network flow.

Theorem

(b) For any inequality condition in (i) or (ii), there exists an incomplete dataset $\hat{\rho} \in \Re_{+}^{\mathcal{M} \setminus \mathcal{M}^{*}}$ that violates the condition but satisfies all the other inequality conditions in (i) and (ii).

Comments:

- Not only does it gives a necessary and sufficient condition, but also it is minimal in the sense of (b).
- It is known that the set of random utility model is a polytope. Our conditions specify all facet defining inequalities of the polytope.
- In general, it is known that obtaining a tight necessary and sufficient condition is very hard.

Implication

- In empirical IO, people often put all unobservable alternatives together and treat them as one outside option, even when the analyst knows which elements are in X^*.
- Theorem implies that this approach may ignore some features of random utility model; more precisely, it does not consider conditions (ii) in Theorem 1.
- Our contribution is to demonstrate this difference clearly by providing a minimal set of testable conditions for observed choice probabilities to be consistent with a random utility model.

In the paper.....

- We prove the result using a network flow theory.
  - To test random utility model, network flow approach is more efficient than the standard linear programming approach because of network flow structure.
- We obtained a bound for missing choice frequencies.
- We have an application to a real dataset

Sketch of the proof

- Our problem is to find the following:

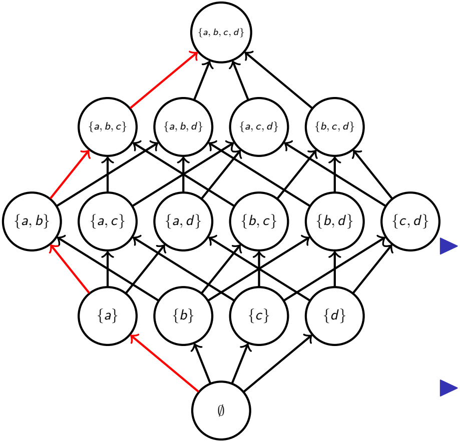
$$(p1) \quad \mu \in \Delta(\mathcal{L})$$

such that for any $(D, x) \in \mathcal{M} \setminus \mathcal{M}^*$,

$$\rho(D, x) = \mu(\succ \in \mathcal{L} \mid x \succ y \text{ for all } y \in D \setminus \{x\}).$$

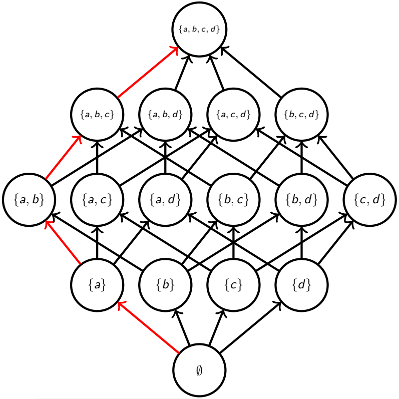
- We rewrite the problem into an existence of a flow in a network.

P2– network flow (Fiorini, 2004)



- Fiorini proved Falmagne's result by using network flow theory:
  - Each node is a subset of X
  - Each arc is connecting a node D and $D \cup x$.
  - For each ranking there is a directed path from $\emptyset$ to X
- We can view random utility models as distributions over paths.
- Each path corresponding a ranking $\succ$ is assigned a flow $\mu(\{\succ\})$.

P2 – network flow (Fiorini, 2004)



◂ One to One mapping

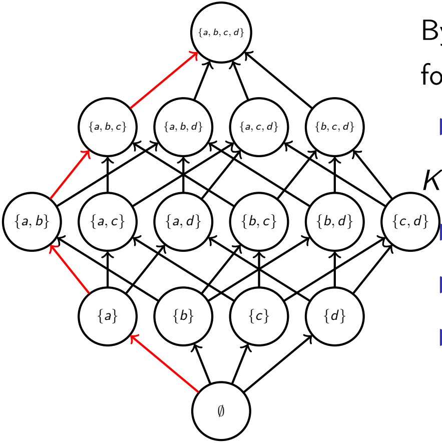
By the construction, we have following restrictions on flows:

- Flow $r(D, D \cup x)$ of an arc from D to $D \cup x$ equals to

$\mu(\{\succ | D^c \succ x \succsim D\}) = K(\rho, D, x)$.

- (Any flow of an arc) ≥ 0.
- (Inflow) $=$ (outflow)
- (The Sum of flows into X) $= 1$.

P2 – network flow (Fiorini, 2004)



By the construction, we have following restrictions:

- Flow of an arc from $D$to$D \cup x$ is $K(\rho, D, x) = \mu(\{\succ | D^c \succ x \succsim D\})$.
- (Any flow of an arc) ≥ 0.
- (Inflow) $=$ (outflow)
- (The Sum of flows into X) $= 1$.

Rewriting the problem

We showed that (P1) is equivalent to the following:

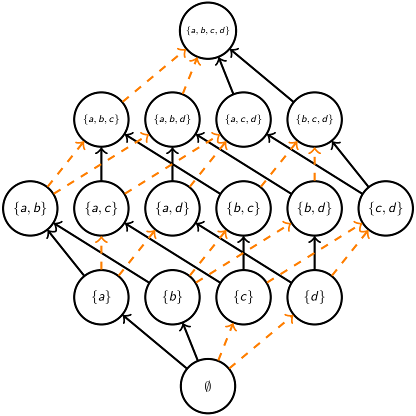
$$(P2) \quad r \in \Re^{\{(D \setminus x, D) | (D,x) \in \mathcal{M}\}}$$

such that

$$\begin{cases} r(D \setminus x, D) = K(\rho, D, x) \text{ for all } (D, x) \in \mathcal{M} \setminus \mathcal{M}^*, \\ \quad \text{(All observable flows are determined by BS polynomials)} \\ r(D \setminus x, D) \geq 0, \\ \quad \text{(All flows are nonnegative)} \\ \sum_{x \in D} r(D \setminus x, D) = \sum_{y \notin D} r(D, D \cup y) \text{ for all } D \in D, \\ \quad \text{(Inflows to } D) = \text{(Outflows from } D) \\ \sum_{x \in X} r(X \setminus x, X) = 1. \\ \quad \text{(The sum of flows into } X \text{ is one).} \end{cases}$$

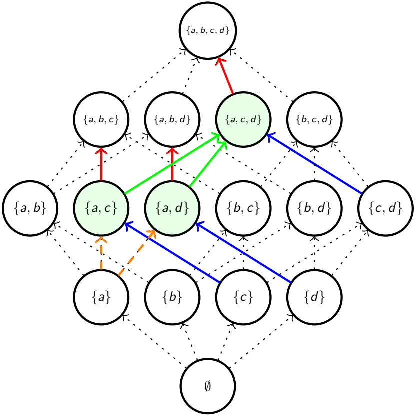
Our approach

- Let $X = \{a, b, c, d\}$ and $X^* = \{c, d\}$
- Given the observable data, we can calculate flows of some arcs (black flows).
- The problem becomes: under what conditions on observable flows we can fill in all unknown flows (yellow flows) in a way that satisfies all restrictions?



Our solution

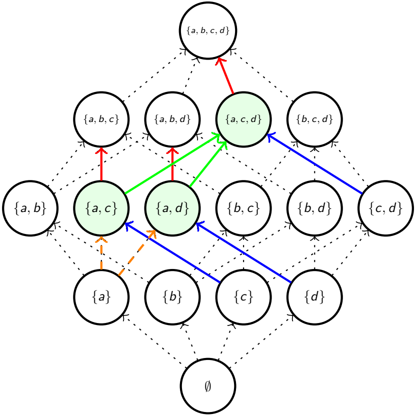
- For each test collection, we check that the sum of inflows must be equal to the sum of outflows.
- For the test collection $\mathcal{C} = \{\{a,c\},\{a,d\},\{a,c,d\}\}$, red are outflows, blue are inflows, yellow are unknown inflows.



  - Note that there are no unobservable outflows by the defintion of test collections.

Our solution

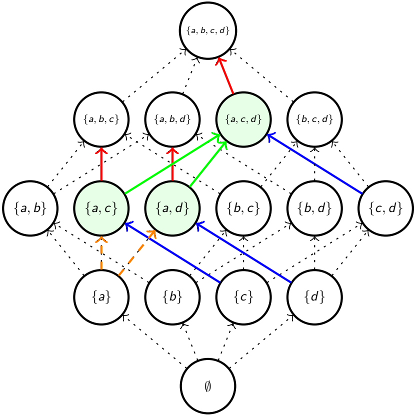
- Let $\mathcal{C} = \{\{a, c\}, \{a, d\}, \{a, c, d\}\}$.
- (Red out-flows)= (Yellow in-flows) + (Blue in-flows)
- Since (Yellow in-flows) ≥ 0, we have (Red out-flows) $-$ (Blue in-flows) ≥ 0!



- Our conditions: These inequalities must be non-negative for all partitions.

Our solution

- $\mathcal{C} = \{\{a, c\}, \{a, d\}, \{a, c, d\}\}$.
- (Red out-flows)
  $$= \sum_{(D,x):D\in\mathcal{C},D\cup x\notin\mathcal{C}} K(D\cup x, x)$$
- (Blue in-flows)
  $$= \sum_{(F,y):F\notin\mathcal{C},F\cup y\in\mathcal{C}} K(\rho, F\cup y, y)$$
- Condition (ii) of the threom is equivalent to (Red out-flows) $-$ (Blue in-flows) ≥ 0.



Theorem

(a) An incomplete dataset $\rho \in \Re_{+}^{\mathcal{M}\setminus\mathcal{M}^*}$ is RU-rationalizable if and only if the following two conditions hold:

- (i) for any $(D, x) \in \mathcal{M} \setminus \mathcal{M}^*$ such that $1 < |D| < |X|$, $K(\rho, D, x) \geq 0$,
- (ii) for any essential test collection $\mathcal{C}$,

$$\left(\sum_{(D,x):D\in\mathcal{C},D\cup x\notin\mathcal{C}} K(\rho, D\cup x, x) - \sum_{(F,y):F\notin\mathcal{C},F\cup y\in\mathcal{C}} K(\rho, F\cup y, y) \right) \geq 0$$

Comment:

- The explanation so far is for the necessity of the condition.
- For sufficiency, we proved a *feasibility theorem* of a network.

◂ details

Theorem

(a) An incomplete dataset $\rho \in \Re_{+}^{\mathcal{M}\setminus\mathcal{M}^*}$ is RU-rationalizable if and only if the following two conditions hold:

- (i) for any $(D, x) \in \mathcal{M} \setminus \mathcal{M}^*$ such that $1 < |D| < |X|$, $K(\rho, D, x) \geq 0$,
- (ii) for any essential test collection $\mathcal{C}$,

$$\left(\sum_{(D,x):D\in\mathcal{C},D\cup x\notin\mathcal{C}} K(\rho, D\cup x, x) - \sum_{(F,y):F\notin\mathcal{C},F\cup y\in\mathcal{C}} K(\rho, F\cup y, y) \right) \geq 0$$

Non redundancy

(b) For any inequality condition in (i) or (ii), there exists an incomplete dataset $\hat{\rho} \in \Re_{+}^{\mathcal{M} \setminus \mathcal{M}^*}$ that violates the condition but satisfies all the other inequality conditions in (i) and (ii).

- Remember

$$\delta(\mathcal{C}) \equiv \sum_{(D,x):D \in \mathcal{C}, D \cup x \notin \mathcal{C}} K(\rho, D \cup x, x) - \sum_{(F,y):F \notin \mathcal{C}, F \cup y \in \mathcal{C}} K(\rho, F \cup y, y).$$

- We fix any essential test collection $\mathcal{C}$.
- We show that there exists a flow such that $\delta(\mathcal{C}) < 0$ but $\delta(\mathcal{C}') \geq 0$ for all other test collection $\mathcal{C}'$.

Interim Summary

- Question: What is a necessary and sufficient condition for an incomplete dataset ρ to be consistent with a random utility model?
- We provide the tight necessary and sufficient condition.
- High level take away is:
  - In empirical IO, people often put all unobservable alternatives together and treat them as one outside option, even when the analyst knows which elements are in X^*.
  - Theorem implies that this approach may ignore some features of random utility model; more precisely, it does not consider conditions (ii) in Theorem 1.
- Our approach can be useful for practical purposes such as obtaining bounds for unobservable choice frequencies

Bounds for unobservable choice probabilities

- Remember that the transportation example: the government does not know how people commute unless they use public transportation, i.e., $X = \{\text{bus, train, walk, drive}\}$ and $X^* = \{\text{walk, drive}\}$
- Analyst is often interested in predicting choice probabilities on unobservable menus.
- For instance, that the government is considering introducing a new tax on gasoline to encourage people to commute by public transportation.

Bounds for unobservable choice probabilities

- To assess the potential impact of the new policy, the government would like to estimate the fraction of people who commute by private car.
- We identify the possible upper and lower bounds on the proportion of drivers following the analysis of Manski (2007).
- Let $\rho \in \Re_{+}^{\mathcal{M} \setminus \mathcal{M}^*}$ be a given incomplete dataset.
- Let Γ be the set of complete data $\hat{\rho}$ that is RU-consistent with the given incomplete dataset ρ.

Bounds for unobservable choice probabilities

- Note that by using (P1), Γ can be written as follows:

$$\left\{ \hat{\rho} \in \Re_+^{\{(D,x)|x \in D \in 2^X\}} \;\middle|\; \begin{array}{l} \text{There exists a } \mu \in \Delta(\mathcal{L}) \text{ that solves (P1)} \\ \text{and saisfies } \rho = \hat{\rho} \text{ on } \mathcal{M} \setminus \mathcal{M}^* \end{array} \right\}$$

- Given $(D, x) \in \mathcal{M} \setminus \mathcal{M}^*$, we want to identify the bounds for $\hat{\rho}(D, x)$ for some $\hat{\rho} \in \Gamma$.
- By the equivalence between, (P1) and (P2), we can rewrite the set Γ as follows:

$$\left\{ \hat{\rho} \in \Re_+^{\{(D,x)|x \in D \in 2^X\}} \;\middle|\; \begin{array}{l} \text{There exists a flow } r \text{ that sloves (P2)} \\ \text{and satisfies } \hat{\rho} = \rho \text{ on } \mathcal{M} \setminus \mathcal{M}^* \end{array} \right\}$$

Bounds for unobservable choice probabilities

- As in Manski (2007), the identification region is convex and all the conditions are linear, so that that for each $\rho^*(D, x)$ for some $(D, x) \in \mathcal{M}^*$, is an interval.
- Compared with the identified region in terms of (P1), this formulation (P2) using the network flow has a computational advantage.
- Because of the network structure, the matrix becomes an incident matrix, which allows us to compute the bounds efficiently. (Even if you do not buy the axiomatic characterization, our approach is useful.)

Bounds for unobservable choice probabilities

- Alternative "efficient" but naive approach to obtaining a bound is simply ignore RU rationalizability.
- To see how much we can tight the bounds by considering RU rationalizability formally, we apply this method to a stochastic choice dataset from the experiment conducted by McCausland et al. (2020).

Bounds for unobservable choice probabilities

- In the experiment, the authors fixed a set $X = \{0, 1, 2, 3, 4\}$ of five lotteries and asked 141 participants to choose one from each subset of X
- Each participant made decision six times for each menu.
- We aggregate these choice frequencies to construct a complete dataset ρ. (We modified ρ so that it becomes RU rationalizable.)
- In this exercise, we mask the choice probabilities of lotteries 0 and 1 and pretend not to observe them; in other words, we set $X^* = \{0, 1\}$ and $\mathcal{D} = 2^X$.

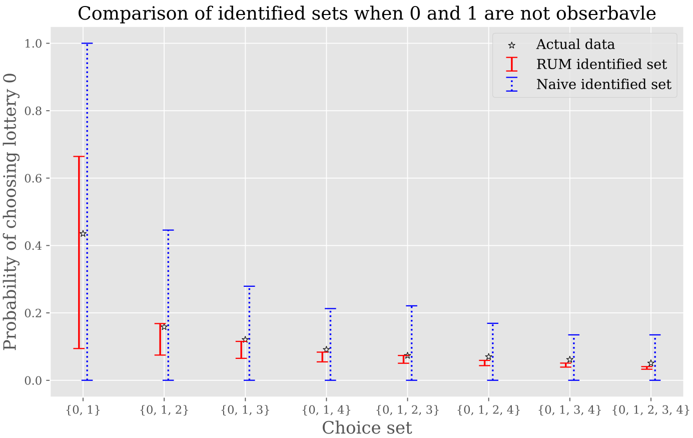
Bounds for unobservable choice probabilities

- Under this setup, we will compute two types of bounds of the probability of lottery 0 being chosen in a given choice set D that contains both lotteries 0 and 1.
- One of them is the trivial bound that is calculated by

$$\left[0, 1 - \sum_{x \in D \cap \{2,3,4\}} \rho(D, x)\right].$$

- The other one is the bound that takes the RU-rationalizabity into account and is computed by the linear program (P2).
- The goal here is to examine how much the random utility assumption shrinks the identified set and improves the prediction of unobservable choice probabilities.

Bounds for unobservable choice probabilities



- Overall, the identified sets of the random utility model, shown in red, are much smaller than the naive bounds, shown in blue, especially when the choice set is large.

Relationship with Mcfadden and Richter (1991)

- Mcfadden and Richter (1991) provide a characterization of the random utility model.
- Unlike the characterization of Falmagne (1978), the characterization of Mcfadden and Richter (1991) holds even for the case when the dataset is incomplete.
- On the other hand, the conditions of Mcfadden and Richter (1991) involve infinite number of sequences and some of the condition are redundant.

Mcfadden and Richter (1991)

- Let $\mathcal{F}$ be the set of pairs (D, x) such that $\rho(D, x)$ is observable to the analysts. $\mathcal{P}(\mathcal{F})$ denotes the set of stochastic choice functions on $\mathcal{F}$.
- Remember that $\mathcal{M} = \{(D, x) \in \mathcal{D} \times X | x \in D\}$ and $\mathcal{M}^* = \{(D, x) \in 2^X \times X | x \in X^* \text{ or } D \notin \mathcal{D}\}$.
- We assume $\mathcal{F} = \mathcal{M} \setminus \mathcal{M}^*$.
- In Falmagne, $\mathcal{F} = \{(D, x) \in 2^X \times X | x \in D\}$. In Mcfadden and Richter, $\mathcal{F}$ is any subset of $\mathcal{M}$.

Definition 2

(Mcfadden and Richter polynomias)Let $\mathcal{F} \subseteq \mathcal{M}$ and $\rho \in \mathcal{P}(\mathcal{F})$. For any sequence $(D_i, x_i)_{i=1}^n$ in $\mathcal{M}$ define

$$R((D_i, x_i)_{i=1}^n, \rho) = \max_{\succ \in \mathcal{L}} \sum_{i=1}^n 1\{x_i \succ D_i \setminus x_i\} - \sum_{i=1}^n \rho(D_i, x_i).$$

Mcfadden and Richter (1991)

Theorem 3

(Mcfadden and Richter(1991))Let $\mathcal{F} \subseteq \mathcal{M}$ and $\rho \in \mathcal{P}(\mathcal{F})$. For any $\rho \in \mathcal{P}(\mathcal{F})$, ρ is a random utility function if and only if $R((D_i, x_i)_{i=1}^n, \rho) \geq 0$ for any sequence $(D_i, x_i)_{i=1}^n$ in $\mathcal{F}$.

- Notice that same (D, x) appears arbitrary many times in the sequence $(D_i, x_i)_{i=1}^n$.
- Although there are finitely many pairs (D, x), the number of the sequences to be tested is infinite.

Mcfadden and Richter (1991)

Definition 4

(i) For any positive integer m, a sequence $(D_i, x_i)_{i=1}^n$ in $\mathcal{M}$ is said to be of repetition of upto m if for each $(D, x) \in \mathcal{M}$ such that $|D| > 2$,

$$\#\{i \mid (D_i, x_i) = (D, x)\} \leq m.$$

(ii) a sequence $(D_i, x_i)_{i=1}^n$ in $\mathcal{M}$ is called *redundant* if $\exists D \in \{D_i\}_{i=1}^n$ such that $\forall x \in D, \exists i$ such that $(D, x) = (D_i, x_i)$.

Proposition

- When the dataset is complete (i.e., $\mathcal{F} = \{(D, x) | x \in D \in 2^X\}$), we can show that considering non-redundant sequences that of repetition of upto 2 is enough to characterize the random utility model.
  - A BS polynomial can be written as a MR polynomial for non-redundant sequences of repetition of up to 2.
  - However, the converse does not hold: Not all MR polynomials for non-redundant sequences of repetition of up to 2 is a BS polynomial.

Proposition

- When the dataset is incomplete (in our sense), we can show that we need check more sequences.
    - By our theorem, we can calculate exactly how many repetitions of non-redundant sequences we need.
    - However, even such an improvement of Mcfadden and Richter's (1991) result would involve redundancy, unlike our theorem.
    - This is because not all MR polynomials for non-redundant sequences correspond to a BS polynomial

Lemma 5

A nonnegative solution ρ^* exists to $(P1)$ if and only if a nonnegative solution r^* exists to (P2).

Proof:

- Define $f : \mathbb{R}^{\mathcal{M}} \to \mathbb{R}^{\{(D\setminus x, D) | (D,x) \in \mathcal{M}\}}$ such that

$$f(p)(D \setminus x, D) = K(p, D, x)$$

- Given the solution ρ^* of P1, $f(\rho^*)$ becomes a solution of P2.
- Inverse of $f : \mathbb{R}^{\{(D\setminus x, D) | (D,x) \in \mathcal{M}\}} \to \mathbb{R}^{\mathcal{M}}$ exists

$$f^{-1}(r)(D, x) = \sum_{E : E \supseteq D} r(E \setminus x, E)$$

- Given solution r^* of P2, $f^{-1}(r^*)$ becomes a solution of P1.

◂ Back

Lemma

- A collection $\mathcal{C} \in 2^{X^*}$ is said to be complete in X^*

$$D \in \mathcal{C} \implies \forall x \in X^*, D \cup x \in \mathcal{C}.$$

- **Lemma:** A solution to (P2) exists if and only if $\delta(\mathcal{C}) \geq 0$ for any complete collection $\mathcal{C}$ in X^*, where

$$\delta(\mathcal{C}) = \left(\sum_{(D,x):D\in\mathcal{C},D\cup x\notin\mathcal{C}} K(\rho, D \cup x, x) - \sum_{(E,y):E\notin\mathcal{C},E\cup y\in\mathcal{C}} K(\rho, E \cup y, y) \right)$$
$$+1\{t \in \mathcal{C}, s \notin \mathcal{C}\} - 1\{s \in \mathcal{C}, t \notin \mathcal{C}\}.$$

  - If the collection is not complete, then
    - there is an unobservable outflow, thus there is no testable implication.
    - u becomes infinity; thus the inequality holds.

Lemma

- Remember a test collection $\mathcal{C} \subset 2^X$ is such that

  $$\mathcal{C} = \{A \cup E \mid E \in \mathcal{E}\},$$

  where $\mathcal{E}$ is a upper set of X^* in 2^{X^*} and $A \subseteq X \setminus X^*$.

- **Lemma** If $\delta(\mathcal{C}) \geq 0$ is for any test collection $\mathcal{C}$, then $\delta(\mathcal{E}) \geq 0$ for any complete collection $\mathcal{E}$ in X^*.
  - Step 1: $\delta(\mathcal{C}) = \sum_{D \in \mathcal{C}} \delta(D)$
  - Step 2: For any $A \subseteq X \setminus X^*$, let $\mathcal{C}_A \equiv \{D \in \mathcal{C} | D \setminus X^* = A\}$. Notice
  - $\mathcal{C}_A \cap \mathcal{C}_B = \emptyset$ if $A \cap B = \emptyset$
  - $\mathcal{C} = \bigcup_{A \subseteq X \setminus X^*} \mathcal{C}_A$
  - $\delta(C) = \sum_{A \subseteq X \setminus X^*} \delta(\mathcal{C}_A)$.

Lemma

- Remember a test collection $\mathcal{C} = \{A \cup E \mid E \in \mathcal{E}\}$ is essential if
  - $A \neq \emptyset$ and $A \subsetneq X \setminus X^*$,
  - $\mathcal{E} \neq 2^{X^*}$.
- **Lemma:** No need to check non-essential test collections.

◂ Details

- Combining results, we have:
  - A solution to (P2) exists if and only if $K(\rho, D, x) \geq 0$ for all $(D, x) \in \mathcal{M}$ such that $x \notin X^*$, and $\delta(\mathcal{C}) \geq 0$ for any essential collection $\mathcal{C}$.

◂ Back

Observations

- If $\mathcal{C} = \{A \cup E \mid E \in 2^{X^*}\}$ for some $A \subset X \setminus X^*$, then $\delta(\mathcal{C}) = 0$.
- If $\mathcal{C} = \{(X \setminus X^*) \cup E | E \in \mathcal{E}\}$ for some $\mathcal{E} \subset 2^{X^*}$ then $\delta(\mathcal{C}) \geq 0$.
- If $\mathcal{C} = \{\emptyset \cup E | E \in \mathcal{E}\}$ for some $\mathcal{E} \subset 2^{X^*}$ then $\delta(\mathcal{C}) \geq 0$.

◂ Back

Network flow

- *Network* is a pair of a node set $\mathcal{N}$ and a set of edges $\mathcal{A} \subset \mathcal{N} \times \mathcal{N}$.
- A function $f : \mathcal{A} \to \Re$ is called a *flow* on a network $(\mathcal{N}, \mathcal{A})$.
- Two nodes s (source) and t (terminal) play special roles. All flows are from s; all flows are into t.
- In our setup,

$$
\begin{aligned}
&\mathcal{N} = 2^X, \\
&\mathcal{A} = \{(D, D \cup x) | D \subset X, x \notin D\}, \\
&s = \emptyset, \\
&t = X.
\end{aligned}
$$

Feasibility Theorem

Let $x \in \mathcal{N}$, $f(x, \mathcal{N}) \equiv \sum_{y \in \mathcal{N}} f(x,y)$; $f(\mathcal{N}, x) \equiv \sum_{y \in \mathcal{N}} f(y,x)$.

Let $l, u : \mathcal{A} \to \Re_+$ such that $l(x,y) \leq u(x,y)$ for any $(x,y) \in \mathcal{A}$.

There exists $f : \mathcal{A} \to \Re_+$ such that

$$
\begin{align}
& f(x, \mathcal{N}) - f(\mathcal{N}, x) = 0 \quad \forall x \in N \setminus \{s, t\}, \tag{1} \\
& f(\mathcal{N}, t) = 1, \tag{2} \\
& l(x,y) \leq f(x,y) \leq u(x,y) \quad \forall (x,y) \in \mathcal{A} \tag{3}
\end{align}
$$

if and only if the following condition holds for all $\mathcal{C} \subset \mathcal{N}$

$$
\sum_{(x,y) \in \mathcal{C} \times \mathcal{C}^c} u(x,y) - \sum_{(x,y) \in \mathcal{C}^c \times \mathcal{C}} l(x,y) \geq \begin{cases} 1 & \text{if } t \notin \mathcal{C}, s \in \mathcal{C}, \\ -1 & \text{if } t \in \mathcal{C}, s \notin \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases}
$$

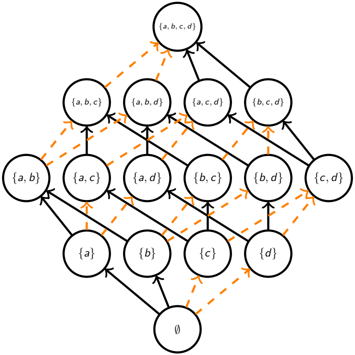
Feasibility Theorem

- We prove the result by using the min-cut maximum-flow theorem.
- Notice that the left hand side of the inequality is the upper bound of flow coming out of $\mathcal{C}$ minus the lower bound of flow coming into $\mathcal{C}$. On the other hand, the right hand side is the actual net outflow at $\mathcal{C}$.

$$\sum_{(x,y)\in\mathcal{C}\times\mathcal{C}^c} u(x,y) - \sum_{(x,y)\in\mathcal{C}^c\times\mathcal{C}} l(x,y) \geq \begin{cases} 1 & \text{if } t \notin \mathcal{C}, s \in \mathcal{C}, \\ -1 & \text{if } t \in \mathcal{C}, s \notin \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases}$$

- We apply the theorem to the network flow defined by (P2).

Application to P2



- To apply the lemma to P2, let

$$l(D \setminus x, D) = K(\rho, D, x) \text{ if } x \notin X^*,$$
$$u(D \setminus x, D) = K(\rho, D, x) \text{ if } x \notin X^*,$$
$$l(D \setminus x, D) = 0 \text{ if } x \in X^*,$$
$$u(D \setminus x, D) = +\infty \text{ if } x \in X^*.$$

Feasibility Theorem

There exists $f : \mathcal{A} \to \Re_+$ such that

$$\begin{aligned} & f(x, \mathcal{N}) - f(\mathcal{N}, x) = 0 \quad \forall x \in N \setminus \{s, t\}, \\ & f(\mathcal{N}, t) = 1, \\ & l(x, y) \leq f(x, y) \leq u(x, y) \quad \forall (x, y) \in \mathcal{A} \end{aligned}$$

if and only if the following condition holds for all $\mathcal{C} \subset \mathcal{N}$

$$\sum_{(x,y) \in \mathcal{C} \times \mathcal{C}^c} u(x, y) - \sum_{(x,y) \in \mathcal{C}^c \times \mathcal{C}} l(x, y) \geq \begin{cases} 1 & \text{if } t \notin \mathcal{C}, s \in \mathcal{C}, \\ -1 & \text{if } t \in \mathcal{C}, s \notin \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases}$$

Feasibility Theorem

- The feasibility theorem requires to test any test collection of nodes.
- Need lemmas to show checking only essential test collections is sufficient. ◂ Details
- Also for any test collection $\mathcal{C}$,

$$\sum_{(x,y)\in\mathcal{C}\times\mathcal{C}^c} u(x,y) - \sum_{(x,y)\in\mathcal{C}^c\times\mathcal{C}} l(x,y)$$

$$= \sum_{(D,x):D\in\mathcal{C},D\cup x\notin\mathcal{C}} K(\rho, D\cup x, x) - \sum_{(F,y):F\notin\mathcal{C},F\cup y\in\mathcal{C}} K(\rho, F\cup y, y).$$

Feasibility Theorem

- The feasibility theorem requires to test any test collection of nodes.
- Need lemmas to show checking only essential test collections is sufficient.
- Also for any test collection $\mathcal{C}$,

$$\sum_{(x,y)\in\mathcal{C}\times\mathcal{C}^c} u(x,y) - \sum_{(x,y)\in\mathcal{C}^c\times\mathcal{C}} l(x,y)$$
$$= \sum_{(D,x):D\in\mathcal{C},D\cup x\notin\mathcal{C}} K(\rho, D\cup x, x) - \sum_{(F,y):F\notin\mathcal{C},F\cup y\in\mathcal{C}} K(\rho, F\cup y, y).$$

◂ back

P1

- Our problem is to find the following:

$$(p1) \quad \mu \in \Delta(\mathcal{L})$$

such that for any $(D, x) \in \mathcal{M} \setminus \mathcal{M}^*$,

$$\rho(D, x) = \mu(\succ \in \mathcal{L} \mid x \succ y \text{ for all } y \in D \setminus \{x\}).$$

◂ Back