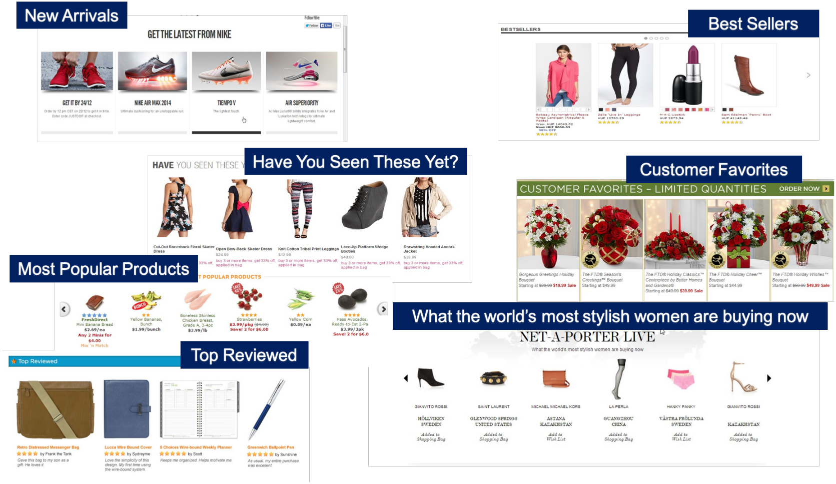
Decision Making with Recommendations

Paul Cheung and Yusufcan Masatlioglu
UT Dallas and University of Maryland

UCLA

Nov 17, 2022

Online Shopping



Endorsements

- “Amazon’s Choice”
- “Superhost”
- “Etsy’s picks”
- “Best Seller”
- “Editor’s pick”
- ...

Others

Basic Observation

- Recommending a product increases the sales of recommended products
    - Senecal and Nantel (2004): Wine/Calculators
    - Gupta and Harris (2010): Computer
    - Adomavicius et al (2018): Digital Music
    - Kawaguchi et al. (2019): Vending machine
    - Farronato et al. (2020): Home services
    - Rietveld et al. (2021): Microloans
    - Bairathi et al. (2022): Freelance
    - ...

Basic Observation

- Recommending a product increases the sales of recommended products
  - Senecal and Nantel (2004): Wine/Calculators
  - Gupta and Harris (2010): Computer
  - Adomavicius et al (2018): Digital Music
  - Kawaguchi et al. (2019): Vending machine
  - Farronato et al. (2020): Home services
  - Rietveld et al. (2021): Microloans
  - Bairathi et al. (2022): Freelance
  - ...
- But HOW?

Two Channels

- Recommendation enlarges awareness set of consumer
    - Recommendation signage: Best Seller, Award Winner (e.g. Goodman et al, 2013)
    - (electronic) Word-of-mouth (e.g. Gupta and Harris, 2010)
    - Uninformative advertising (e.g. Mayzlin and Shin, 2011)
- Recommendation affects consumer's valuation
    - Consumer's Rating (Cosley et al, 2003)
    - Willingness to Pay (Adomavicius et al, 2018)
    - Consumer's utility (Kawaguchi et al, 2021)

A Puzzle

- Does Recommendation affect choices through attention or preferences?

A Puzzle

- Does Recommendation affect choices through attention or preferences?
  - Informational: enlarging awareness set of consumer
  - Persuasive: increasing consumer's evaluation

Aim

- To understand how choice is affected by recommendation

Aim

- To understand how choice is affected by recommendation
- Distinguish different channels of recommendation from observed choices

Aim

- To understand how choice is affected by recommendation
- Distinguish different channels of recommendation from observed choices
- Provide a new theoretical foundation for applied and empirical studies on recommendation

How do we proceed?

- Deterministic

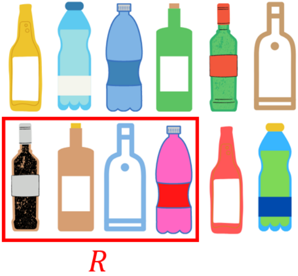
How do we proceed?

- Deterministic
- Probabilistic

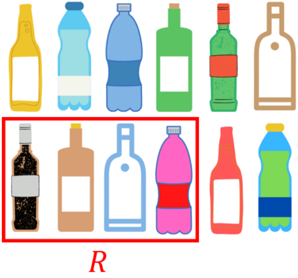
How do we proceed?

- Deterministic
- Probabilistic
  - Non-Parametric
  - Parametric

Decision Problem

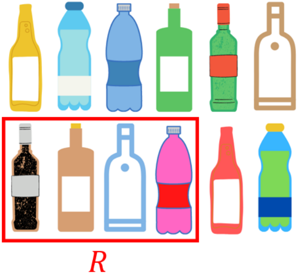


Decision Problem



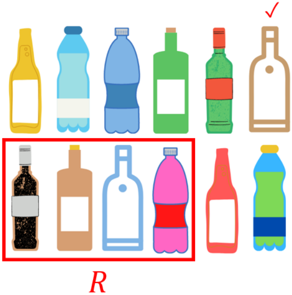
- X: set of alternatives
  - a dataset consisting of a single and fixed menu
  - variation comes from different recommendation sets

Decision Problem



- X: set of alternatives
  - a dataset consisting of a single and fixed menu
  - variation comes from different recommendation sets
- $c : 2^X \to X$, a choice function
- $c(R) \in X$ for $R \subseteq X$

Decision Problem



- In standard model, $c(S) \in S$
- Here, $c(R) \in X$

How to model the choice

$$c(R) = ???$$

How to model the choice

$$c(R) = \max(X, \succ)$$

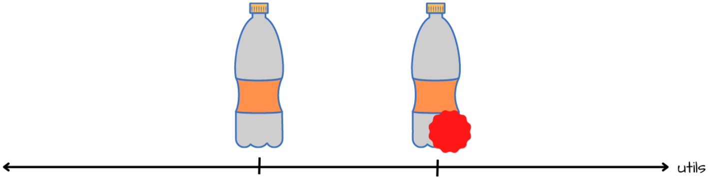
How to model the choice

$$c(R) = \max(X, \succ)$$

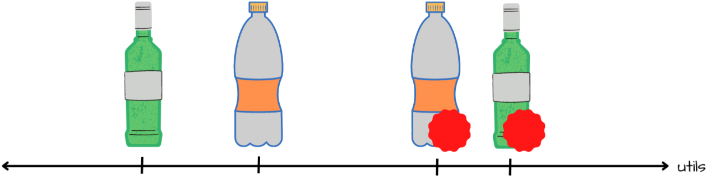
- Assuming *away* the effect of recommendation
- How to model??

Persuasive Recommendation (PR)

Persuasive Recommendation (PR)

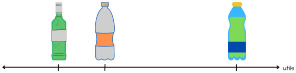


Persuasive Recommendation (PR)



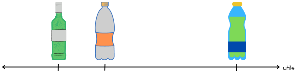
Persuasive Recommendation (PR)

- $\succ$ - Preference on X

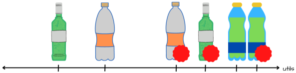


Persuasive Recommendation (PR)

- $\succ$ - Preference on X

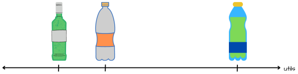


- $\succ^{\star}$ - Preference on $X \cup X^{*}$

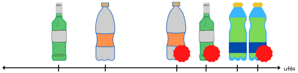


Persuasive Recommendation (PR)

- $\succ$ - Preference on X



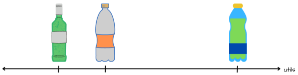
- $\succ^\star$ - Preference on $X \cup X^*$



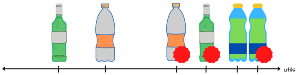
- The relationship between $\succ$ and $\succ^\star$
  - $x^* \succ^\star x$
  - $x \succ^\star y \Leftrightarrow x \succ y$

Persuasive Recommendation (PR)

- $\succ$ - Preference on X



- $\succ^\star$ - Preference on $X \cup X^*$



- The relationship between $\succ$ and $\succ^\star$
  - $x^* \succ^\star x$
  - $x \succ^\star y \Leftrightarrow x \succ y$
- Persuasive Recommendation Model (PR)

$$c(R) = \max{}^*(R^* \cup X, \succ^\star)$$

Informational Recommendation (IR)

Informational Recommendation (IR)

Limited Attention

Informational Recommendation (IR)

Limited Attention

Informational Recommendation (IR)

Limited Attention

Informational Recommendation (IR)

- $\succ$ - Preference on X

Informational Recommendation (IR)

- $\succ$ - Preference on X
- Assume limited consideration
- a denotes the best option in her consideration set

Informational Recommendation (IR)

- $\succ$ - Preference on X
- Assume limited consideration
- a denotes the best option in her consideration set
- Informational Recommendation Model (IR)

$$c(R) = \max(R \cup a, \succ)$$

Informational Recommendation (IR)

- $\succ$ - Preference on X
- Assume limited consideration
- a denotes the best option in her consideration set
- Informational Recommendation Model (IR)

$$c(R) = \max(R \cup a, \succ)$$

- Equivalently,

$$c(R) = \max(R \cup A, \succ) \text{ where } a = \max(A, \succ)$$

- A: awareness set

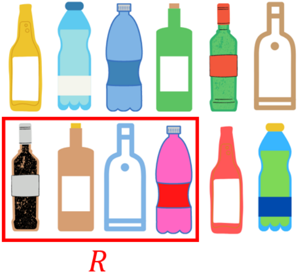
Behavioral implication

- Suppose we do not know which one is the correct model

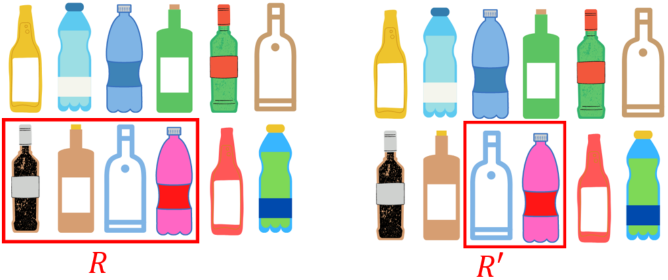
Behavioral implication

- Suppose we do not know which one is the correct model
    - How do we distinguish them from observed choice behavior?

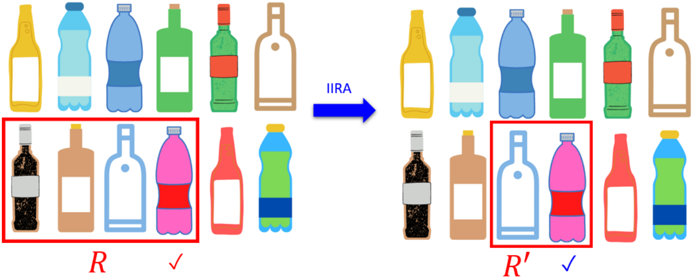
Independence of Irrelevant *Recommended* Alternatives (IIRA)



Independence of Irrelevant *Recommended* Alternatives (IIRA)

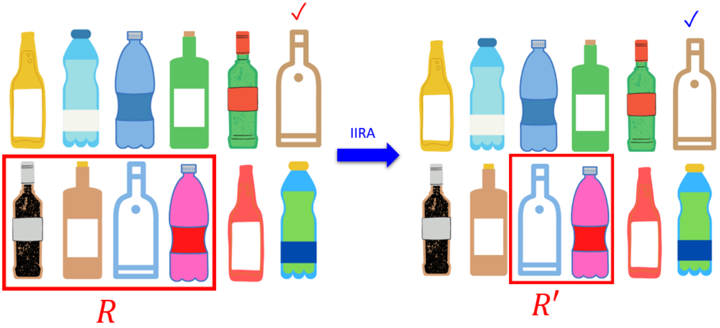


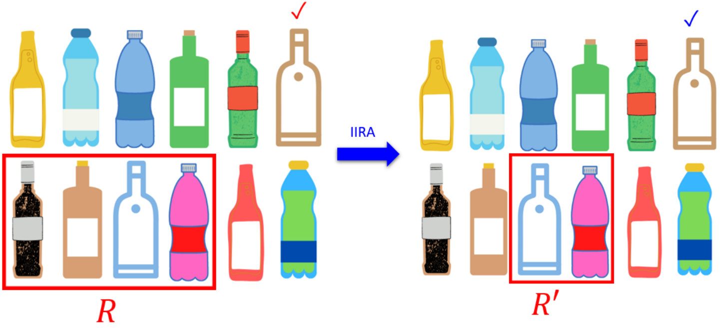
Independence of Irrelevant *Recommended* Alternatives (IIRA)



If $R' \subseteq R$ and $c(R) \in R'$, then $c(R) = c(R')$

Independence of Irrelevant *Recommended* Alternatives (IIRA)

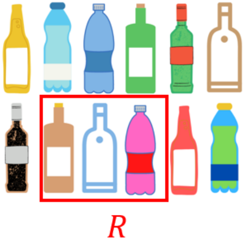




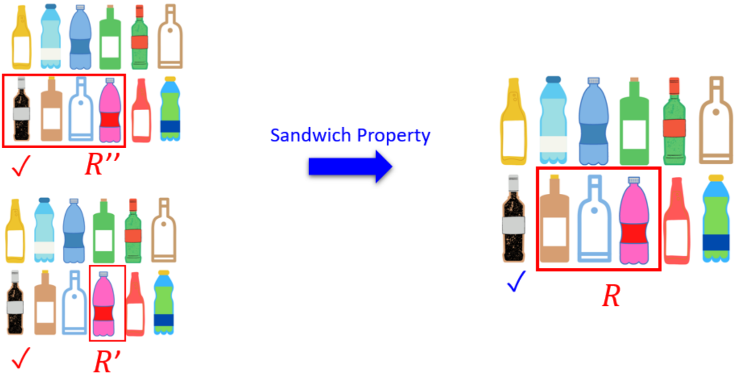
IIRA

If $R' \subseteq R$ and $c(R) \notin R \setminus R'$, then $c(R) = c(R')$

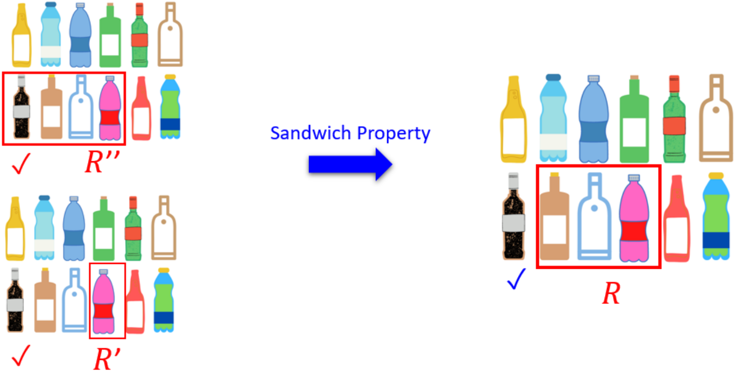

✓
R''
✓
R'


R

Sandwich Property



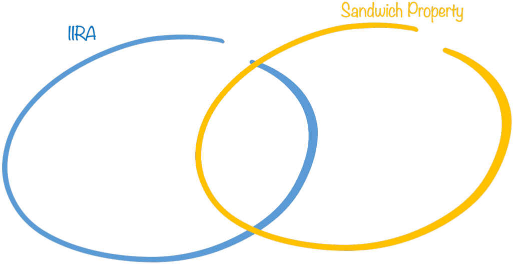
Sandwich Property



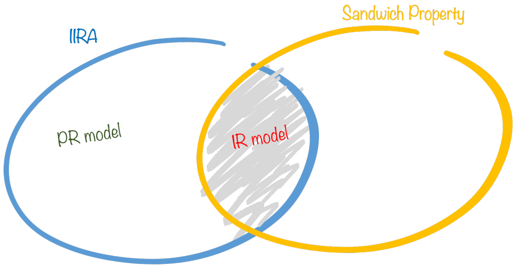
Sandwich Property

If $R' \subseteq R \subseteq R''$ and $c(R'') = c(R')$, then $c(R) = c(R')$.

Axioms



Axioms



Let $\mathcal{D}$ includes all recommendation sets with $|R| \leq 3$.

Theorem (Preference Channel)

c has a PR representation on $\mathcal{D}$ if and only if c satisfies Axiom IIRA.

Theorem (Attention Channel)

c has a IR representation on $\mathcal{D}$ if and only if c satisfies Axiom IIRA, and Sandwich Property.

Identification for Preference Channel

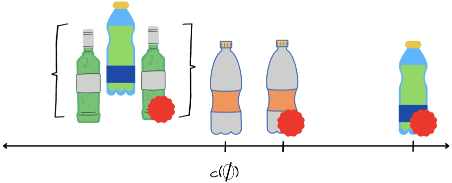
- Let c belong to the PR Model

Identification for Preference Channel

- Let c belong to the PR Model
- Revelations on $\succ^*$ from choices

Identification for Preference Channel

- Let c belong to the PR Model
- Revelations on $\succ^*$ from choices
  - $c(\emptyset)$ must be the best alternative in $\succ$
  - Preference over the upper contour set of $c(\emptyset)$ is identified
  - Preference over the lower contour set of $c(\emptyset)$ is NOT identified



Identification for Attention Channel

- Let c belong to the IR Model

Identification for Attention Channel

- Let c belong to the IR Model
- Revelations on $(a, \succ)$ from choices

Identification for Attention Channel

- Let c belong to the IR Model
- Revelations on $(a, \succ)$ from choices
  - $c(\emptyset)$ must be the default option
  - Preference over the upper contour set of $c(\emptyset)$ is identified
  - Preference over the lower contour set of $c(\emptyset)$ is NOT identified

Up until here

- Provided a new theoretical framework for recommendation in the deterministic environment
- Discovered when we can distinguish utility channel from attention channel
- Identify the primitives of the models

Probabilistic Data

Probabilistic Data

- Real-world data often comes in the form of probabilistic choice
    - Aggregate data
    - Repeated choice

Probabilistic Data

- Real-world data often comes in the form of probabilistic choice
  - Aggregate data
  - Repeated choice
- Building based on deterministic models
  - Non-parametric models (à la RUM)
  - Parametric models (à la Luce)

Probabilistic Data

- $\rho(x, R)$: frequency of x being chosen when R is the recommended set
- A single and fixed menu (no menu variation)
  - variation comes from different recommendation sets
- $\sum_{x \in X} \rho(x, R) = 1$

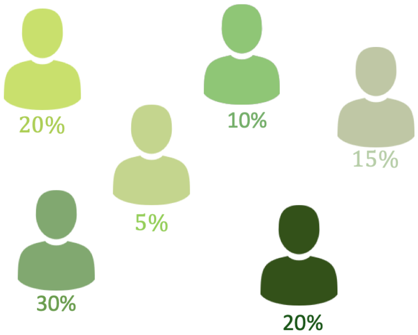
Non-Parametric Model

Heterogeneous Population

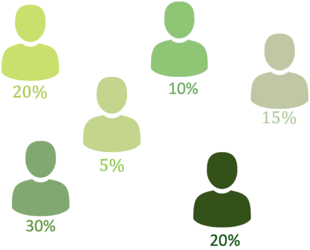
Think of a stochastic choice coming from a heterogeneous population

Heterogeneous Population

Think of a stochastic choice coming from a heterogeneous population



Classical Random Utility Model (RUM)



- μ: a probability distribution over all preference types

$$\rho(x, R) = \sum_{x \text{ is the best in } \succ} \mu(\succ)$$

- the randomness in choices is attributed to the variation in tastes or types

PR-RUM (Preference Channel)

- Each type is denoted by $\succ^\star$ (as in the PR Model)
- μ: probability measure over the set of all $\succ^\star$ on $X \cup X^*$

$$\rho(x, R) = \sum_{\substack{x \text{ is chosen} \\ \text{by type } \succ^\star}} \mu(\succ^\star)$$

PR-RUM (Preference Channel)

- Each type is denoted by $\succ^\star$ (as in the PR Model)
- μ: probability measure over the set of all $\succ^\star$ on $X \cup X^*$

$$\rho(x, R) = \sum_{\substack{x \text{ is chosen} \\ \text{by type } \succ^\star}} \mu(\succ^\star)$$

- Rich type space
    - If $n = 3$, 90 types in PR-RUM vs 6 types in RUM
    - If $n = 4$, 2520 types in PR-RUM vs 24 types in RUM

IR-RUM (Attention Channel)

IR-RUM (Attention Channel)

- Each type is denoted by $(a, \succ)$ (as in the IR Model)
- μ: probability measure over the set of all $(a, \succ)$ where $\succ$ on X

$$\rho(x, R) = \sum_{\substack{x \text{ is chosen} \\ \text{by type } (a,\succ)}} \mu(a, \succ)$$

IR-RUM (Attention Channel)

- Each type is denoted by $(a, \succ)$ (as in the IR Model)
- μ: probability measure over the set of all $(a, \succ)$ where $\succ$ on X

$$\rho(x, R) = \sum_{\substack{x \text{ is chosen} \\ \text{by type } (a,\succ)}} \mu(a, \succ)$$

- Rich type space
  - If $n = 3$, 18 types in IR-RUM vs 6 types in RUM
  - If $n = 4$, 96 types in IR-RUM vs 24 types in RUM

Behavioral Implications of RUM

- Remember RUM
- Has preference maximization any implications for aggregate data?

Behavioral Implications of RUM

- Remember RUM
- Has preference maximization any implications for aggregate data?
- For RUM in the standard environment,
  - The Block-Marschak polynomials are non-negative.
  - Given ρ,

$$q_\rho(x,S) := \sum_{B \supseteq S} (-1)^{|B \setminus S|} \rho(x,B) \geq 0$$

- Choice data can be represented by RUM iff $q_\rho(x,R) \geq 0$.

- Given ρ,

$$q_\rho(x,S) := \sum_{B \supseteq S} (-1)^{|B \setminus S|} \rho(x,B) \geq 0$$

Behavioral Implications of RUM

- Given ρ,

$$q_\rho(x, S) := \sum_{B \supseteq S} (-1)^{|B \setminus S|} \rho(x, B) \geq 0$$

- $q_\rho(x, S)$: the probability of types who rank x behind the elements of $X \setminus S$ and ahead of the elements in S

- Given ρ,

$$q_\rho(x,S) := \sum_{B \supseteq S} (-1)^{|B \setminus S|} \rho(x,B) \geq 0$$

- $q_\rho(x,S)$: the probability of types who rank x behind the elements of $X \setminus S$ and ahead of the elements in S
- For example, if $X = \{a,b,c,d\}$, then $q(b,\{b,c\})$ identifies the probability of $a \succ d \succ b \succ c$ and $d \succ a \succ b \succ c$

Behavioral Implications of RUM

- Given ρ,

$$q_\rho(x, S) := \sum_{B \supseteq S} (-1)^{|B \setminus S|} \rho(x, B) \geq 0$$

- $q_\rho(x, S)$: the probability of types who rank x behind the elements of $X \setminus S$ and ahead of the elements in S

- For example, if $X = \{a, b, c, d\}$, then $q(b, \{b, c\})$ identifies the probability of $a \succ d \succ b \succ c$ and $d \succ a \succ b \succ c$

- Hence $\mu(\{a \succ d \succ b \succ c, d \succ a \succ b \succ c\}) = q(b, \{b, c\})$

Block-Marschak Polynomials

For $x \in R$,

$$q_\rho(x, R) := \sum_{B \supseteq R} (-1)^{|B \setminus R|} \rho(x, B)$$

Block-Marschak Polynomials

For $x \in R$,

$$q_\rho(x,R) := \sum_{B \supseteq R} (-1)^{|B \setminus R|} \rho(x,B)$$

A new object

For $x \notin R$,

$$y_\rho(x,R) := \sum_{x \notin B \supseteq R} (-1)^{|B \setminus R|} \rho(x,B)$$

Behavioral Implication

Non-negativity of BM

For $a \in R$, $q(a, R) \geq 0$ and $y(a, R \setminus a) \geq 0$.

Behavioral Implication

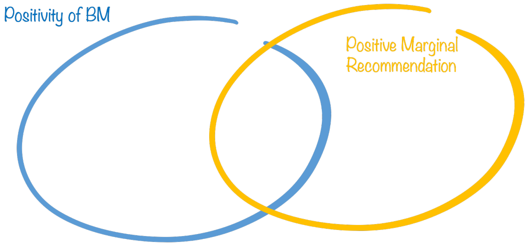
Non-negativity of BM

For $a \in R$, $q(a, R) \geq 0$ and $y(a, R \setminus a) \geq 0$.

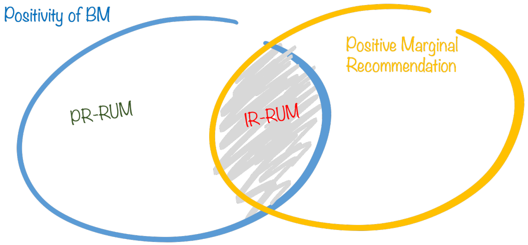
Positive Marginal Recommendation

For $a \in R$, $q(a, R) \geq y(a, R \setminus a)$.

Axioms



Axioms


Positivity of BM
Positive Marginal Recommendation
PR-RUM
IR-RUM

Characterization

Assume $\mathcal{D} = 2^X$

Theorem

ρ is a PR-RUM if and only if ρ satisfies Non-negativity of BM.

Theorem

ρ is an IR-RUM if and only if ρ satisfies Non-negativity of BM and Positive Marginal Recommendation.

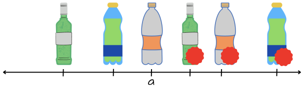
Proof

Identification for Preference Channel

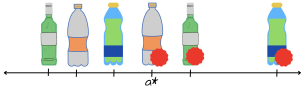
- Revelations on μ defined over $\succ^*$ from choices

Identification for Preference Channel

- Revelations on μ defined over $\succ^*$ from choices
  - $y_\rho(a, \emptyset)$: the probability of types who rank a as the best alternative in $\succ$ and $b^* \succ^* a$ for all b



  - $q_\rho(a, \{a\})$: the probability of types who rank a^* above b for all and below b^* for all b



Identification for Attention Channel

- Revelations on μ defined over $(a, \succ)$ from choices

Identification for Attention Channel

- Revelations on μ defined over $(a, \succ)$ from choices
  - $y_\rho(b, A)$: the probability of types who rank b just above A and b is their default
  - $q_\rho(b, A \cup \{b\})$: the probability of types who rank b just above A and their default is within $A \cup \{b\}$

Note that $y_\rho(a, \emptyset) = q_\rho(a, \{a\})$

So far

- Distinguish utility channel from attention channel in probabilistic world
- Identification of types
- No parametric assumptions...

Parametric

Parametric

- tractable
- strong uniqueness properties
- sharp identification results for application purposes

The MNL (Luce) Model

- Most used parametric model
- Specifies a utility $u(x)$ for each alternative x
- Probability of choosing an alternative x in a set X

$$\frac{u(x)}{\sum_{y \in X} u(y)}$$

- We now apply this idea to a model with recommendations

PR-Luce

- $u'(x)$: the utility of x with recommendation
- $u(x)$: the utility of x w/o recommendation
- $u'(x) \geq u(x)$: positive recommendation
- When x is recommended

$$\frac{u'(x)}{\sum_{y \in R} u'(y) + \sum_{y \in X \setminus R} u(y)}$$

PR-Luce

- $u'(x)$: the utility of x with recommendation
- $u(x)$: the utility of x w/o recommendation
- $u'(x) \geq u(x)$: positive recommendation
- When x is recommended

$$\frac{u'(x)}{\sum_{y \in R} u'(y) + \sum_{y \in X \setminus R} u(y)}$$

- When x is not recommended

$$\frac{u(x)}{\sum_{y \in R} u'(y) + \sum_{y \in X \setminus R} u(y)}$$

PR-Luce

Positivity: $\rho(x, R) > 0$ even if $x \notin R$

PR-Luce

Positivity: $\rho(x, R) > 0$ even if $x \notin R$

Definition

A choice rule ρ has a persuasive Luce recommendation representation (PR-Luce) if there exists functions $u, u' : X \to \mathbb{R}_{++}$ such that for $x \in X$, $u'(x) \geq u(x)$ and

$$\rho^{PR}(x, R) := \begin{cases} \frac{u'(x)}{\sum_{y \in R} u'(y) + \sum_{y \in X \setminus R} u(y)} & \text{if } x \in R \\ \frac{u(x)}{\sum_{y \in R} u'(y) + \sum_{y \in X \setminus R} u(y)} & \text{otherwise} \end{cases}$$

for all $R \in \mathcal{D}$.

- Alternatively, we can write $u'(x) = u(x)r(x)$
  - $r(x) \geq 1$ captures the increase in weight for alternative x

IR-Luce

Fix the default: a,

$$\text{Probability being chosen } := \begin{cases} \frac{u(x)}{\sum_{z \in R \cup a} u(z)} & \text{if } x \in R \cup a \\ 0 & \text{otherwise} \end{cases}$$

Fix the default: a,

$$\text{Probability being chosen } := \begin{cases} \frac{u(x)}{\sum_{z \in R \cup a} u(z)} & \text{if } x \in R \cup a \\ 0 & \text{otherwise} \end{cases}$$

and

$$d(a): \text{ probability of } a \text{ is being the default}$$

IR-Luce

Definition

A choice rule ρ has a informational Luce recommendation representation (IR-Luce) if there exists functions $u : X \to \mathbb{R}_{++}$ and $d : X \to \mathbb{R}_{++}$ with $\sum_{x \in X} d(x) = 1$ such that

$$\rho^{IR}(x, R) = \begin{cases} \sum_{z \in X} d(z) \frac{u(x)}{\sum_{y \in R \cup z} u(y)} & \text{if } x \in R \\ d(x) \frac{u(x)}{\sum_{y \in R \cup x} u(y)} & \text{otherwise} \end{cases}$$

for $x \in X$ and $R \in \mathcal{D}$.

Axioms

Axioms

Axiom: Recommended Luce-IIA

For $x, y \in R \cap R'$,

$$\frac{\rho(x, R)}{\rho(y, R)} = \frac{\rho(x, R')}{\rho(y, R')}$$

Axiom: R-Path Independence

For $x \notin R$ and $x \cup R \subseteq R'$,

$$\rho(x, R)\rho(x \cup R, R') \text{ is independent of } R$$

Axioms

Axiom: R-Regularity

For $x \notin R$, $\rho(x, R) \leq \rho(x, R \setminus y)$.

- It is implied by R-Path Independence.

Axioms

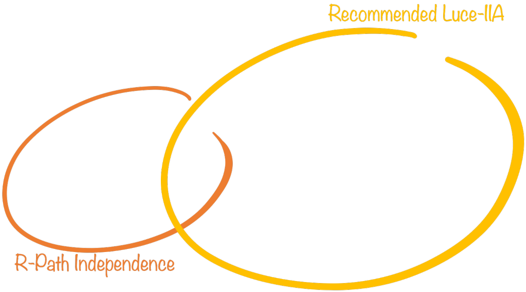
Axiom: Strong Luce-IIA

For $x, y \in R \cap R'$, $t, z \notin R \cup R'$,

$$\frac{\rho(x,R)}{\rho(x,R')} = \frac{\rho(y,R)}{\rho(y,R')} = \frac{\rho(t,R)}{\rho(t,R')} = \frac{\rho(z,R)}{\rho(z,R')}$$

- It (immediately) implies Recommended Luce-IIA.

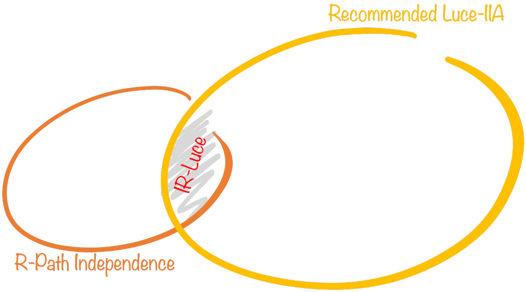
Axioms



R-Path Ind.: For $x \notin R$ and $R \cup x \subseteq R'$, $\rho(x, R)\rho(R \cup x, R')$ is independent of R

Rec. Luce-IIA: For $x, y \in R \cap R'$, $\frac{\rho(x,R)}{\rho(y,R)} = \frac{\rho(x,R')}{\rho(y,R')}$

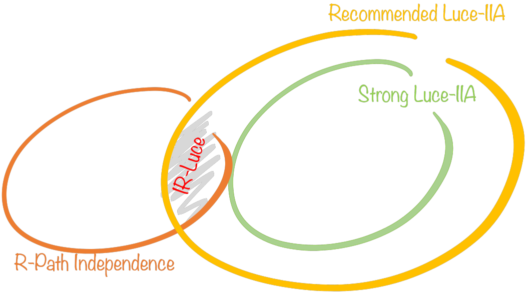
Axioms



R-Path Ind.: For $x \notin R$ and $R \cup x \subseteq R'$, $\rho(x, R)\rho(R \cup x, R')$ is independent of R

Rec. Luce-IIA: For $x, y \in R \cap R'$, $\frac{\rho(x,R)}{\rho(y,R)} = \frac{\rho(x,R')}{\rho(y,R')}$

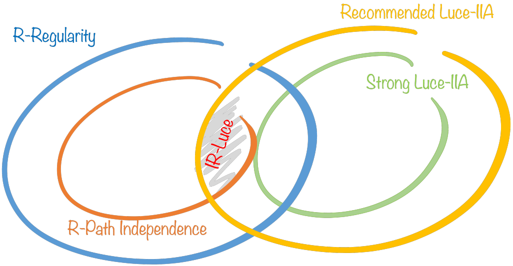
Axioms



R-Path Ind.: For $x \notin R$ and $R \cup x \subseteq R'$, $\rho(x, R)\rho(R \cup x, R')$ is independent of R

Rec. Luce-IIA: For $x, y \in R \cap R'$, $\frac{\rho(x,R)}{\rho(y,R)} = \frac{\rho(x,R')}{\rho(y,R')}$

Strong Luce-IIA: For $x, y \in R \cap R'$, $t, z \notin R \cup R'$, $\frac{\rho(x,R)}{\rho(x,R')} = \frac{\rho(y,R)}{\rho(y,R')} = \frac{\rho(t,R)}{\rho(t,R')} = \frac{\rho(z,R)}{\rho(z,R')}$



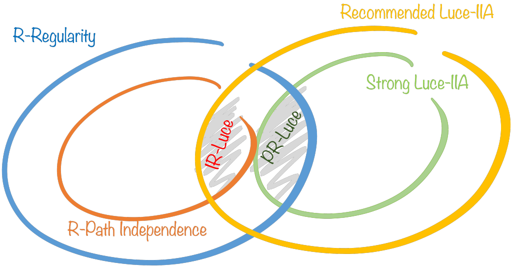
R-Path Ind.: For $x \notin R$ and $R \cup x \subseteq R'$, $\rho(x, R)\rho(R \cup x, R')$ is independent of R

Rec. Luce-IIA: For $x, y \in R \cap R'$, $\frac{\rho(x,R)}{\rho(y,R)} = \frac{\rho(x,R')}{\rho(y,R')}$

Strong Luce-IIA: For $x, y \in R \cap R'$, $t, z \notin R \cup R'$, $\frac{\rho(x,R)}{\rho(x,R')} = \frac{\rho(y,R)}{\rho(y,R')} = \frac{\rho(t,R)}{\rho(t,R')} = \frac{\rho(z,R)}{\rho(z,R')}$

R-Regularity : For $x \notin R$, $\rho(x, R) \leq \rho(x, R \setminus y)$.

Axioms



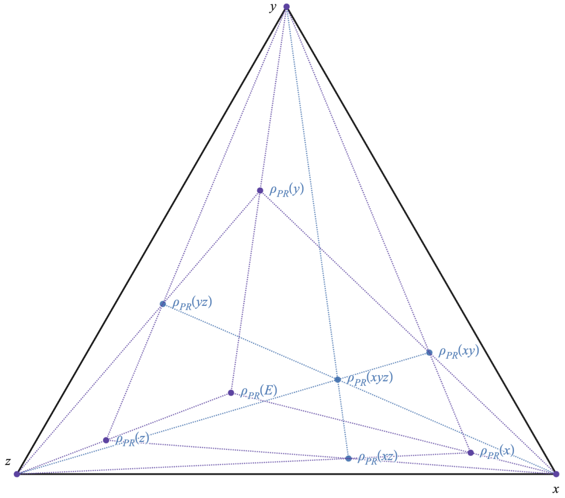
R-Path Ind.: For $x \notin R$ and $R \cup x \subseteq R'$, $\rho(x, R)\rho(R \cup x, R')$ is independent of R

Rec. Luce-IIA: For $x, y \in R \cap R'$, $\frac{\rho(x,R)}{\rho(y,R)} = \frac{\rho(x,R')}{\rho(y,R')}$

Strong Luce-IIA: For $x, y \in R \cap R'$, $t, z \notin R \cup R'$, $\frac{\rho(x,R)}{\rho(x,R')} = \frac{\rho(y,R)}{\rho(y,R')} = \frac{\rho(t,R)}{\rho(t,R')} = \frac{\rho(z,R)}{\rho(z,R')}$

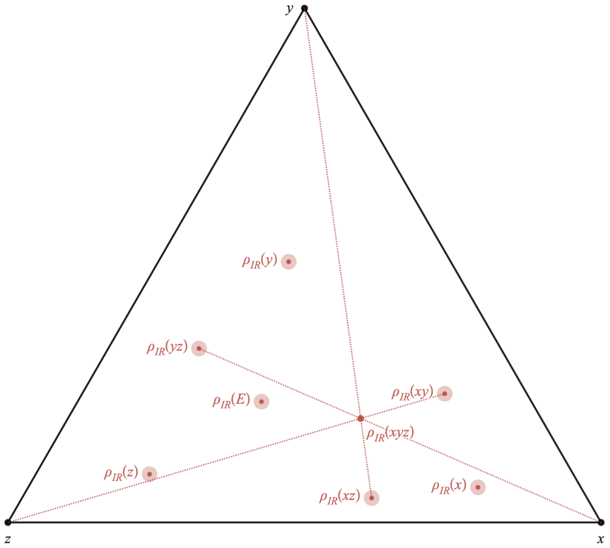
R-Regularity : For $x \notin R$, $\rho(x, R) \leq \rho(x, R \setminus y)$.

PR-LUCE



y
$\rho_{PR}(y)$
$\rho_{PR}(yz)$
$\rho_{PR}(xy)$
$\rho_{PR}(xyz)$
$\rho_{PR}(E)$
$\rho_{PR}(z)$
$\rho_{PR}(xz)$
$\rho_{PR}(x)$
z
x

IR-LUCE



y
$\rho_{IR}(y)$
$\rho_{IR}(yz)$
$\rho_{IR}(E)$
$\rho_{IR}(xy)$
$\rho_{IR}(xyz)$
$\rho_{IR}(z)$
$\rho_{IR}(xz)$
$\rho_{IR}(x)$
z
x

Characterization: Parametric

Theorem (Preference Channel)

Let $\mathcal{D}$ includes all recommendation sets with $|R| \leq 2$. Then, ρ has a PR-Luce representation if and only if ρ satisfies Axiom R-Regularity and General Luce-IIA.

Theorem (Attention Channel)

Assume $\mathcal{D} = 2^X$. Then, ρ has an IR-Luce representation if and only if ρ satisfies Axiom Recommended Luce-IIA and R-Path Independence.

Identification: Parametric

Proposition

Suppose ρ is IR-Luce. Let $\mathcal{D}$ includes recommendation sets $\emptyset$ and $\{a\}$ for some a, then we can fully identify the parameters of the models.

Proposition

Suppose ρ is PR-luce. Let $\mathcal{D}$ includes all recommendation sets with $|R| \leq 1$, then we can fully identify the parameters of the models.

- IR-Luce requires less data to fully identify the parameters.

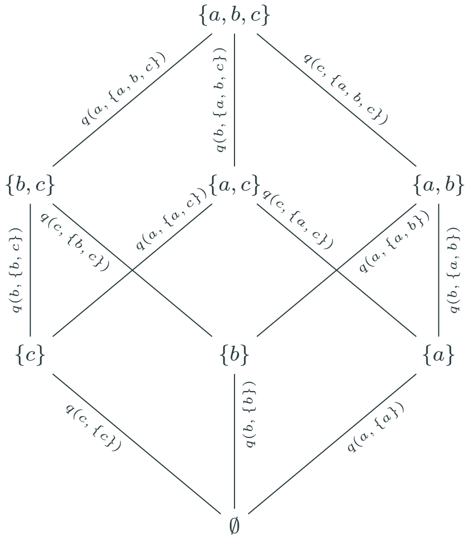
Summary: Parametric

- Provide a framework to study the effect of recommendation
- Distinguish between attention and preference channels
  - Fully distinguish between utility and attention

To conclude

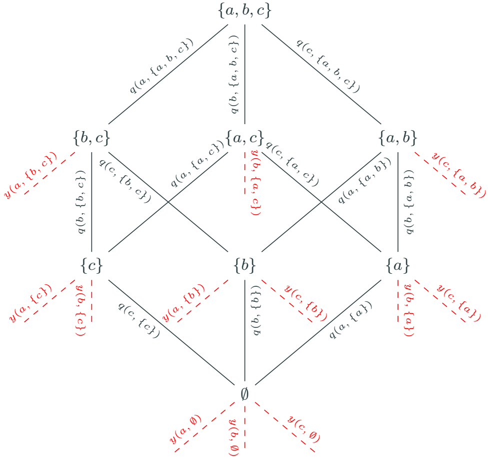
- Provide a framework to study the effect of recommendation
- Distinguish between attention channel and preference channel
- Characterize probabilistic choice models for real-world application
- More to come
    - e.g. Choice effects, Spillover effects, Bounded Rationality

Ideas of Proof



Number of sinking paths: $3! = 6$

Hasse diagram



Number of outgoing paths: $\sum_{k=1}^{3} C_k^3 * k * k! = 33$

Number of outgoing paths with $q(x, R) \rightarrow y(x, R \setminus x)$: $\sum_{k=1}^{3} C_k^3 * k! = 15$

Appendix: Luce's Axiom back

In the standard Environment, the followings are equivalent

Luce's IIA

$$\frac{\rho(x, R)}{\rho(y, R)} = \frac{\rho(x, R')}{\rho(y, R')}$$

Luce's Choice Axiom

$\rho(a, R)\rho(R, R')$ is independent of R

Appendix: Luce's Axiom back

In the standard Environment, the followings are equivalent

Luce's IIA

$$\frac{\rho(x,R)}{\rho(y,R)} = \frac{\rho(x,R')}{\rho(y,R')}$$

Luce's Choice Axiom

$\rho(a,R)\rho(R,R')$ is independent of R

Here, we apply one on recommended, and one on non-recommended.

Recommended IIA

For $x, y \in R \cap R'$,

$$\frac{\rho(x,R)}{\rho(y,R)} = \frac{\rho(x,R')}{\rho(y,R')}$$

R-Path Independence

For $x \notin R$ and $R \cup x \subseteq R'$,

$\rho(x,R)\rho(R \cup x, R')$ is independent of R

Stronger Axiom 6* back

Due to Recommended IIA, for some recommendation set A that includes x and z, we let

$$r(z, x) := \frac{\rho(z, A)}{\rho(x, A)}$$

Stronger Axiom 6* back

Due to Recommended IIA, for some recommendation set A that includes x and z, we let

$$r(z,x) := \frac{\rho(z,A)}{\rho(x,A)}$$

Axiom: 6* (Off-recommendation Independence). For $x \notin R$,

$$\rho(x,R)(1+\sum_{z\in R} r(z,x)) \text{ is independent of } R$$

Discrete Choice perspective

Random utility is defined as

$$U(x) = v(x) + \epsilon(x)$$

where $\epsilon(x)$ is known as "random utility shock".

Discrete Choice perspective

Random utility is defined as

$$U(x) = v(x) + \epsilon(x)$$

where $\epsilon(x)$ is known as "random utility shock".

Let a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Event where x achieves the highest utility in a set A,

$$\omega_{x,A} = \{\omega \in \Omega : U(x) \geq U(y) \text{ for all } y \in A\}$$

R-logit

A choice rule ρ has a R-logit representation if there exists $v : X \to \mathbb{R}$, $d : X \to \mathbb{R}_+$ with $\sum_{x \in X} d(x) = 1$ and $\epsilon : \Omega \to \mathbb{R}^X$ which follows Gumbel distribution with noise parameter λ and is i.i.d. across $x \in X$ such that

$$\rho(x, R) = \begin{cases} \sum_{a \in X} d(a) \mathbb{P}(\omega_{x, R \cup a}) & \text{if } x \in R \\ d(x) \mathbb{P}(\omega_{x,R}) & \text{if } x \notin R \end{cases}$$

R-logit

A choice rule ρ has a R-logit representation if there exists $v : X \to \mathbb{R}$, $d : X \to \mathbb{R}_+$ with $\sum_{x \in X} d(x) = 1$ and $\epsilon : \Omega \to \mathbb{R}^X$ which follows Gumbel distribution with noise parameter λ and is i.i.d. across $x \in X$ such that

$$\rho(x, R) = \begin{cases} \sum_{a \in X} d(a)\mathbb{P}(\omega_{x,R\cup a}) & \text{if } x \in R \\ d(x)\mathbb{P}(\omega_{x,R}) & \text{if } x \notin R \end{cases}$$

Remark. The closed-form solution for $\mathbb{P}(\omega_{x,A})$ is,

$$\mathbb{P}(\omega_{x,A}) = \frac{e^{v(x)/\lambda}}{\sum_{z \in A} e^{v(z)/\lambda}}$$