Mathematical Modeling of Irrational Behavior: towards a linear formulation

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

May 16, 2023

Outline

1. Choice models
2. Beyond rationality
3. Optimization

Decision rule

Homo economicus

Rational and narrowly self-interested economic actor who is optimizing her outcome

Utility

$$U_n : \mathcal{C}_n \longrightarrow \mathbb{R} : a \rightsquigarrow U_n(a)$$

- captures the attractiveness of an alternative
- measure that the decision maker wants to optimize

Behavioral assumption

- the decision maker associates a utility with each alternative
- the decision maker is a perfect optimizer
- the alternative with the highest utility is chosen

Random utility model

Random utility

$$U_{in} = V_{in} + \varepsilon_{in} = \beta^T X_{in} + \varepsilon_{in}.$$

Choice model

$$P(i|\mathcal{C}_n) = \Pr(U_{in} \geq U_{jn}, \forall j \in \mathcal{C}_n),$$

Logit model

Assumptions

ε_{in} are i.i.d. EV(0, μ).

Choice model

$$P_n(i|\mathcal{C}_n) = \frac{y_{in} e^{\mu V_{in}}}{\sum_{j=1}^{J} y_{jn} e^{\mu V_{jn}}}.$$

Outline

Beyond rationality

Motivation

- There is evidence that human beings are not necessarily rational in the way assumed by random utility models.
- We first review some experiments that illustrate that (apparent) irrationality.

Example: pain lovers

[Kahneman et al., 1993]

- Short trial: immerse one hand in water at 14° for 60 sec.
- Long trial: immerse the other hand at 14° for 60 sec, then keep the hand in the water 30 sec. longer as the temperature of the water is gradually raised to 15°.
- Outcome: most people prefer the long trial.
- Explanation: duration plays a small role, the peak and the final moments matter.

Example: The Economist

[Ariely, 2008]

Subscription to The Economist

Web only	@ $59
Print only	@ $125
Print and web	@ $125

Example: The Economist

[Ariely, 2008]

Subscription to The Economist

Experiment 1	Experiment 2
Web only @ $59	Web only @ $59
Print only @ $125	
Print and web @ $125	Print and web @ $125

Example: The Economist

[Ariely, 2008]

Subscription to The Economist

	Experiment 1	Experiment 2	
16	Web only @ $59	Web only @ $59	68
0	Print only @ $125		
84	Print and web @ $125	Print and web @ $125	32

The Economist: explanations

- Dominated alternative.
- According to utility maximization, should not affect the choice.
- But it affects the perception, which affects the choice.

Decoy effect

Decoy

High-price, low-value product compared to other items in the choice set.

Behavior

Consumers shift their choice to more expensive items.

Applications

- Travel and tourism. [Josiam and Hobson, 1995]
- Wine lists in restaurants. [Kimes et al., 2012]
- Tobacco treatment. [Rogers et al., 2020]
- Online diamond retail. [Wu and Cosguner, ta]

Example: good or bad wine?

Choose a bottle of wine...

	Experiment 1	Experiment 2
1	McFadden red at $10	McFadden red at $10
2	Nappa red at $12	Nappa red at $12
3		McFadden special reserve pinot noir at $60
	Most would choose 2	Most would choose 1

- Context plays a role on perceptions.
- Here, perceived quality is increased.

Example: live and let die

[Kahneman and Tversky, 1986]

Population of 600 is threatened by a disease.

Two alternative treatments to combat the disease have been proposed.

	Experiment 1 # resp. = 152	Experiment 2 # resp. = 155	
72%	Treatment A: 200 people saved	Treatment C: 400 people die	22%
28%	Treatment B: 600 saved with prob. 1/3 0 saved with prob. 2/3	Treatment D: 0 die with prob. 1/3 600 die with prob. 2/3	78%

Example: to be free

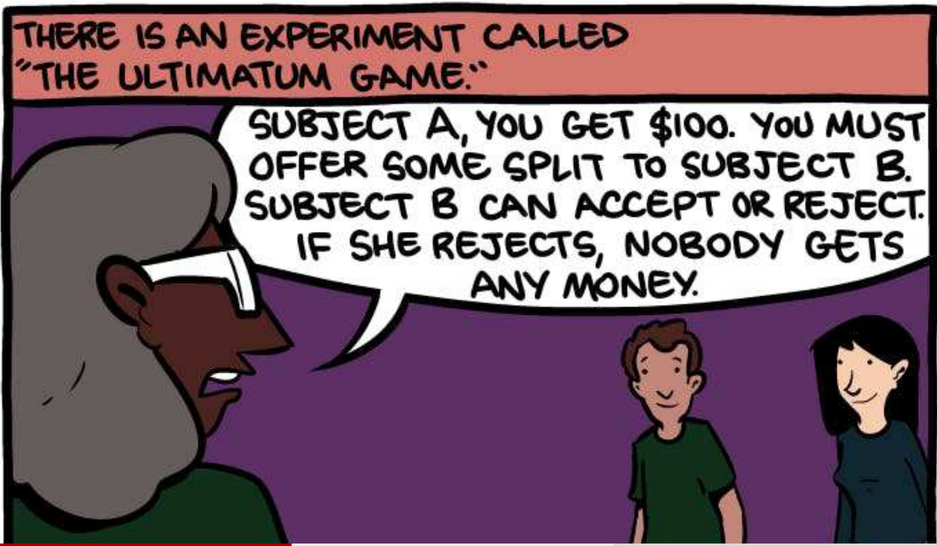
[Ariely, 2008]

Choice between a fine and a regular chocolate

	Experiment 1	Experiment 2
Lindt	\$0.15	\$0.14
Hershey	\$0.01	\$0.00
Lindt chosen	73%	31%
Hershey chosen	27%	69%

Discontinuity at 0

Ultimatum game



Ultimatum game

Ultimatum game

Ultimatum game

Optimal solution

Subject B should accept any offer.

In practice

Offers of less than 30% are often rejected.

Modeling latent concepts

Motivation

- Some observed behavior may appear irrational, and inconsistent with random utility.
- It is only apparent, as these behaviors can be explained by more complex formulations of the concept of utility.
- In particular, this may involve subjective and latent concepts such as perceptions and attitudes.
- Latent concepts can be introduced in choice models.

Indirect measurements of latent concepts

Attitude towards the environment

For each question, response on a scale: strongly agree, agree, neutral, disagree, strongly disagree, no idea.

- The price of oil should be increased to reduce congestion and pollution.
- More public transportation is necessary, even if it means additional taxes.
- Ecology is a threat to minorities and small companies.
- People and employment are more important than the environment.
- I feel concerned by the global warming.
- Decisions must be taken to reduce the greenhouse gas emission.

Indirect measurements of latent concepts

Psychometric indicators

- Usually easy to respond.
- Arbitrary units.
- Important to minimize framing.

Data

For each individual, we have

- Vector of independent variables: x.
- Choice: i.
- vector of psychometric indicators: I.

Prediction model

Latent variable

- Captures perceptions, attitudes, anchors, etc.
- Not observed.
- Modeled as a function of observed variables:

$$X^* = \text{EnvironmentalAttidude} = \text{f}(\text{Age, Education, etc.}; \theta) + \xi.$$

Random utility model

- Utility is also unobserved.
- Modeled as a function of observed variables, as well as the latent variable(s):

$$\text{Utility(PublicTransport)} = \text{f}(\text{Price, Time, Frequency, EnvironmentalAttitude}; \theta) + \varepsilon$$

Prediction model

Choice model: mixture of logit models

$$P_n(i|x_n, X_n^*, \mathcal{C}_n) = \frac{y_{in} e^{\mu V_{in}(x_n, X_n^*)}}{\sum_{j=1}^{J} y_{jn} e^{\mu V_{jn}(x_n, X_n^*)}}.$$

$$\begin{aligned} P_n(i|x_n, \mathcal{C}_n) &= \int_t P_n(i|x_n, t, \mathcal{C}_n) f_{X_n^*}(t) dt \\ &= \int_t \frac{y_{in} e^{\mu V_{in}(x_n, t)}}{\sum_{j=1}^{J} y_{jn} e^{\mu V_{jn}(x_n, t)}} f_{X_n^*}(t) dt. \end{aligned}$$

Outline

Demand-based optimization

Context

- An operator providing goods or services.
- Potentially, competing operators.
- Customers who freely decide which service/good to choose.

Objective

Help the operator with strategic, tactical or operational decisions.

Comments

- This is the core business of operations research.
- But the decisions of customers are often assumed to be given, exogenous.
- Challenge: use choice models to capture the demand, the decisions of customers.

Demand-based optimization

Examples

- Pricing, toll setting.
- Revenue management.
- Facility location.
- Assortment optimization.
- Passenger-centric railway timetabling.
- ...

Main issue

Demand representation

- $d_i(x)$: number of customers who select service/good i, under decision x.
- Using a choice model:

$$d_i(x) = \sum_n P_n(i|\mathcal{C}_n) = \sum_n \int_t \frac{y_{in}(x)e^{\mu V_{in}(x,t)}}{\sum_{j=1}^J y_{jn}(x)e^{\mu V_{jn}(x,t)}} f_{X^*}(t)dt.$$

Issue

- Most optimization models in OR rely on convenient relaxations of the original problem.
- Usually, "convenient" means linear or convex.
- But mixtures of logit models are far from being convex.

Exogenous and endogenous variables

Endogenous variables

- Decision variables of the operator that influence the choice of customers.
- Examples: price, quality of service, properties of goods, etc.

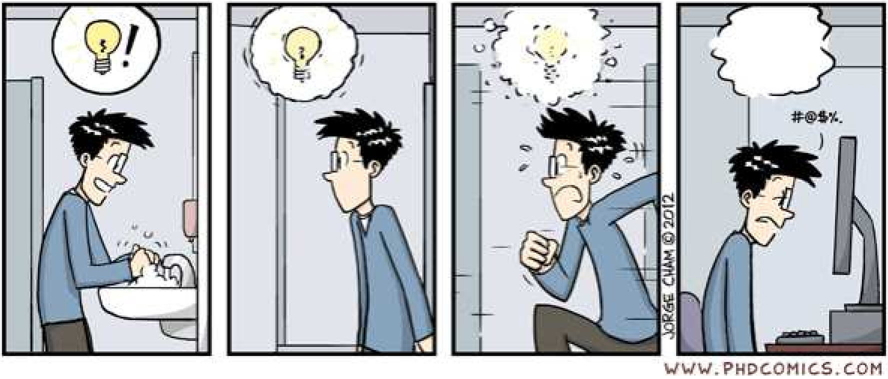
Exogenous variables

- Variables influencing the choice of customers, but not decided by the operator.
- Examples: decisions of the competing operators, attitudes, perceptions, etc.

Mathematical requirement

We need linearity (or convexity) in the endogenous variables.

The main idea



The main idea

Linearization

- Hopeless to linearize the logit formula (we tried...)
- Anyway, we want to go beyond logit.

Idea

Work with the utility and not the probability.

A linear formulation

Latent variable

$$X_n^* = f_X(z_{\text{endo}}, z_{\text{exo}}) + \xi_n, \text{where } f_X \text{ is linear (or convex) in } z_{\text{endo}}.$$

Simulation

- Assume a distribution for ξ_n
- E.g. normal distribution.
- Draw R realizations ξ_{nr}, $r = 1, \ldots, R$

A linear formulation

Utility function

$$U_{in} = V_{in}(x_{\text{endo}}, x_{\text{exo}}, X_n^*) + \varepsilon_{in},$$

where V_{in} is linear (or convex) in x_{endo} and X_n^* (and so, in z_{endo}).

Simulation

- Assume a distribution for ε_{in}
- E.g. logit: i.i.d. extreme value
- Draw R realizations ε_{inr}, $r = 1, \ldots, R$

Scenarios

Draws

- Draw R realizations ξ_{inr}, $r = 1, \ldots, R$
- We obtain R scenarios

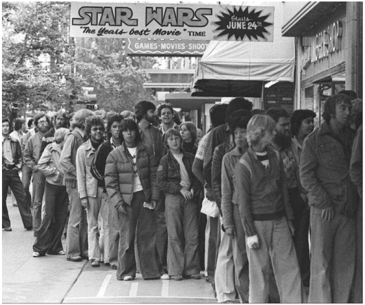
$$X^*_{nr} = \sum_k \theta_k z_{\text{endo}} + f(z_{\text{exo}}) + \xi_{inr}.$$

$$U_{inr} = \sum_k \beta_k x_{\text{endo}} + f(x_{\text{exo}}) + \varepsilon_{inr}.$$

- For each scenario r, we can identify the largest utility.
- It corresponds to the chosen alternative.

Capacities

- Demand may exceed supply
- Each alternative i can be chosen by maximum c_i individuals.
- An exogenous priority list is available.
- Can be randomly generated, or according to some rules.
- The numbering of individuals is consistent with their priority.



Choice set

Variables

$y_i \in \{0,1\}$ operator decision

$y^d_{in} \in \{0,1\}$ customer decision (data)

$y_{in} \in \{0,1\}$ product of decisions

$y_{inr} \in \{0,1\}$ capacity restrictions

Constraints

$$y_{in} = y^d_{in} y_i \quad \forall i, n$$
$$y_{inr} \leq y_{in} \quad \forall i, n, r$$

Utility

Variables

U_{inr}	utility
$z_{inr} = \begin{cases} U_{inr} & \text{if } y_{inr} = 1 \\ \ell_{nr} & \text{if } y_{inr} = 0 \end{cases}$	discounted utility
(ℓ_{nr} smallest lower bound)	

Constraint: utility

$$U_{inr} = \overbrace{\sum_k \beta_k x_{kn,\text{endo}} + f(x_{n,\text{exo}})}^{V_{in}} + \varepsilon_{inr} \; \forall i, n, r$$

Utility (ctd)

Constraints: discounted utility

$$\begin{aligned} \ell_{nr} &\leq z_{inr} && \forall i, n, r \\ z_{inr} &\leq \ell_{nr} + M_{inr} y_{inr} && \forall i, n, r \\ U_{inr} - M_{inr}(1 - y_{inr}) &\leq z_{inr} && \forall i, n, r \\ z_{inr} &\leq U_{inr} && \forall i, n, r \end{aligned}$$

Choice

Variables

$$U_{nr} = \max_{i \in \mathcal{C}} z_{inr}$$

$$w_{inr} = \begin{cases} 1 & \text{if } z_{inr} = U_{nr} \\ 0 & \text{otherwise} \end{cases} \qquad \text{choice}$$

Constraints

$$z_{inr} \leq U_{nr} \qquad \forall i, n, r$$

$$U_{nr} \leq z_{inr} + M_{nr}(1 - w_{inr}) \qquad \forall i, n, r$$

$$\sum_i w_{inr} = 1 \qquad \forall n, r$$

$$w_{inr} \leq y_{inr} \qquad \forall i, n, r$$

Capacity

If capacity reached $\Rightarrow y_{inr} = 0$

$$\sum_{m=1}^{n-1} w_{imr} \leq (c_i - 1)y_{inr} + (n-1)(1 - y_{inr}) \ \forall i > 0, n > c_i, r$$

If capacity not reached $\Rightarrow y_{inr} = 1$

$$c_i(y_{in} - y_{inr}) \leq \sum_{m=1}^{n-1} w_{imr}, \ \forall i > 0, n, r$$

Family of models

Constraints

- Set of linear constraints characterizing choice behavior
- Can be included in any relevant optimization problem.

Examples

- Profit maximization
- Facility location

Difficulties

- big M constraints
- large dimensions

Profit maximization

Profit

If p_{in} is the price paid by individual to purchase option i, the revenue generated by this option is

$$\frac{1}{R}\sum_{r=1}^{R}\sum_{n=1}^{N} p_{in} w_{inr}.$$

Linearization

If $a_{in} \leq p_{in} \leq b_{in}$, we define $\eta_{inr} = p_{in} w_{inr}$, and the following constraints:

$$\begin{aligned} a_{in} w_{inr} &\leq \eta_{inr} \\ \eta_{inr} &\leq b_{in} w_{inr} \\ p_{in} - (1 - w_{inr}) b_{in} &\leq \eta_{inr} \\ \eta_{inr} &\leq p_{in} - (1 - w_{inr}) a_{in} \end{aligned}$$

A case study

Challenge

- Take a choice model from the literature.
- It must be a mixture of logit.
- It must involve heterogeneity.
- Show that it can be integrated in a relevant MILP.

A case study

Challenge

- Take a choice model from the literature.
- It must be a mixture of logit.
- It must involve heterogeneity.
- Show that it can be integrated in a relevant MILP.

Parking choice

- [Ibeas et al., 2014]

Parking choices [Ibeas et al., 2014]

Alternatives

- Paid on-street parking
- Paid underground parking
- Free street parking

Model

- $N = 50$ customers
- $\mathcal{C} = \{\text{PSP}, \text{PUP}, \text{FSP}\}$
- $\mathcal{C}_n = \mathcal{C} \quad \forall n$
- $p_{in} = p_i \quad \forall n$
- Capacity of 20 spots
- Mixture of logit models

General experiments

Uncapacitated vs Capacitated case

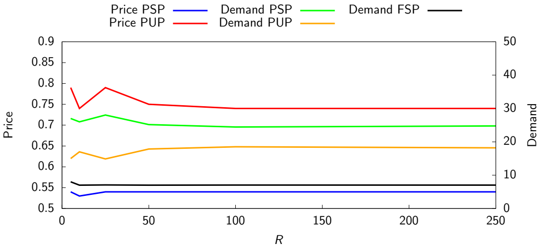
- Maximization of revenue
- Unlimited capacity
- Capacity of 20 spots for PSP and PUP

Price differentiation by population segmentation

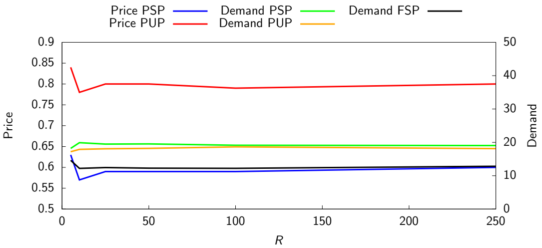
- Reduced price for residents
- Two scenarios
  1. Subsidy offered by the municipality
  2. Operator is forced to offer a reduced price

Uncapacitated vs Capacitated case

Uncapacitated



Capacitated



Computational time

	Uncapacitated case				Capacitated case			
R	Sol time	PSP	PUP	Rev	Sol time	PSP	PUP	Rev
5	2.58 s	0.54	0.79	26.43	12.0 s	0.63	0.84	25.91
10	3.98 s	0.53	0.74	26.36	54.5 s	0.57	0.78	25.31
25	29.2 s	0.54	0.79	26.90	13.8 min	0.59	0.80	25.96
50	4.08 min	0.54	0.75	26.97	50.2 min	0.59	0.80	26.10
100	20.7 min	0.54	0.74	26.90	6.60 h	0.59	0.79	26.03
250	2.51 h	0.54	0.74	26.85	1.74 days	0.60	0.80	25.93

Facility location

Data

- U_{in}: exogenous,
- C_i: fixed cost to open a facility,
- c_i: operational cost per customer to run the facility.

Objective function

$$\min \sum_{i \in \mathcal{C}_k} C_i y_i + \frac{1}{R} \sum_r \sum_i \sum_n c_i w_{inr}$$

Benders decomposition

$$\min \sum_{i \in \mathcal{C}_k} C_i y_i + \frac{1}{R} \sum_r \sum_i \sum_n c_i w_{inr}$$

subject to

$$\max_w U_{nr} = \sum_i U_{inr} w_{inr}$$

$$\sum_i w_{inr} \leq 1$$

$$w_{inr} \leq y_i$$

$$w_{inr} \geq 0$$

$$w_{inr}, y_i \in \{0, 1\}.$$

Benders decomposition

Customer subproblem: fix y_i^*

$$\max_w U_{nr} = \sum_i U_{inr} w_{inr}$$

subject to

$$\sum_i w_{inr} = 1$$

$$w_{inr} \leq y_i^*$$

$$w_{inr} \geq 0.$$

Property

Totally unimodular: no integrality constraint is required.

Benders decomposition

Primal

$$\min_{w} U = -\sum_i U_i w_i$$

subject to

$$\sum_i w_i = 1$$

$$w_i \leq y_i^* \qquad \forall i$$

$$w_i \geq 0.$$

Dual

$$\max_{\lambda,\mu} \lambda + \sum_i \mu_i y_i^*$$

subject to

$$\lambda + \mu_i \leq -U_i \qquad \forall i$$

$$\mu_i \leq 0 \qquad \forall i$$

Bender decomposition

Ongoing work

- Exploit the duality results to generate cuts for the master problem.
- Investigate the use of Benders for other problems.
    - profit maximization,
    - maximum likelihood estimation of the parameters.

Bibliography I

Ariely, D. (2008).
Predictably irrational. The hidden forces that shape our decisions.
Harper Collins.

Ibeas, A., dell'Olio, L., Bordagaray, M., and de D. Ortúzar, J. (2014).
Modelling parking choices considering user heterogeneity.
Transportation Research Part A: Policy and Practice, 70:41 – 49.

Josiam, B. M. and Hobson, J. S. P. (1995).
Consumer choice in context: The decoy effect in travel and tourism.
Journal of Travel Research, 34(1):45–50.

Kahneman, D., Fredrickson, B., Schreiber, C., and Redelmeier, D. (1993).
When more pain is preferred to less: Adding a better end.
Psychological Science, 4(6):401–405.

Bibliography II

Kahneman, D. and Tversky, A. (1986).
Rational choice and the framing of decisions.
Journal of business, 59(4):251–278.

Kimes, S. E., Phillips, R., and Summa, L. (2012).
Pricing in restaurants.
In Özer, O. and Phillips, R., editors, The Oxford Handbook of pricing management, Oxford Handbooks. OUP Oxford.

Rogers, E., Vargas, E., and Voigt, E. (2020).
Exploring the decoy effect to guide tobacco treatment choice: a randomized experiment.
BMC Res Notes, 13(3).

Bibliography III

Wu, C. and Cosguner, K. (t.a.).
Profiting from the decoy effect: A case study of an online diamond retailer.
Marketing Science.