Mathematical Modeling of Human Behavior: application to mobility

Michel Bierlaire

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

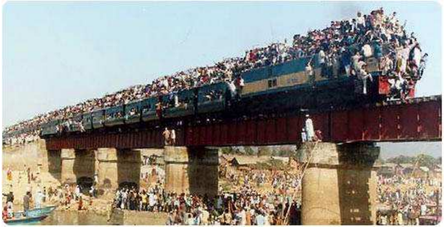
June 14, 2022

Outline

Travel demand models

- Supply = infrastructure
- Demand = behavior, choices
- Congestion = mismatch

Travel demand models



- Usually in OR:
- optimization of the supply
- for a given (fixed) demand

Aggregate demand

- Homogeneous population
- Identical behavior
- Price (P) and quantity (Q)
- Demand functions: $P = f(Q)$
- Inverse demand: $Q = f^{-1}(P)$

TRANSP-OR

EPFL

Disaggregate demand

- Heterogeneous population
- Different behaviors
- Many variables:
    - Attributes: price, travel time, reliability, frequency, etc.
    - Characteristics: age, income, education, etc.
- Complex demand/inverse demand functions.

TRANSP-OR

EPFL

Examples in mobility

Discrete choices

- Choice of activity.
- Choice of destination.
- Choice of mode of transportation.
- Choice of departure time.
- Choice of path.

Outline

Decision rule

Homo economicus

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

Behavioral assumptions

- The decision maker solves an optimization problem.
- The analyst needs to define
  - the decision variables,
  - the objective function,
  - the constraints.

Microeconomic consumer theory

Continuous choice set

- Consumption bundle:

$$q = \begin{pmatrix} q_1 \\ \vdots \\ q_L \end{pmatrix} ; p = \begin{pmatrix} p_1 \\ \vdots \\ p_L \end{pmatrix}$$

- Budget constraint

$$p^T q = \sum_{\ell=1}^{L} p_\ell q_\ell \leq I.$$

Preferences

Operators $\succ$, $\sim$, and $\succsim$

- $q_a \succ q_b$: q_a is preferred to q_b,
- $q_a \sim q_b$: indifference between q_a and q_b,
- $q_a \succsim q_b$: q_a is at least as preferred as q_b.

Rationality

- Completeness: for all bundles a and b,

$$q_a \succ q_b \text{ or } q_a \prec q_b \text{ or } q_a \sim q_b.$$

- Transitivity: for all bundles a, b and c,

$$\text{if } q_a \succsim q_b \text{ and } q_b \succsim q_c \text{ then } q_a \succsim q_c.$$

- "Continuity": if q_a is preferred to q_b and q_c is arbitrarily "close" to q_a, then q_c is preferred to q_b.

Utility

Utility function

- Parameterized function:

$$\widetilde{U} = \widetilde{U}(q_1, \ldots, q_L; \theta) = \widetilde{U}(Q; \theta)$$

- Consistent with the preference indicator:

$$\widetilde{U}(q_a; \theta) \geq \widetilde{U}(q_b; \theta)$$

is equivalent to

$$q_a \succsim q_b.$$

- Unique up to an order-preserving transformation

Optimization

Optimization problem

$$\max_{q} \widetilde{U}(q;\theta)$$

subject to

$$p^T q \leq I, \; q \geq 0.$$

Demand function

- Solution of the optimization problem.
- Quantity as a function of prices and budget.

$$q^* = f(I, p; \theta)$$

Microecomomic theory

How does it work for discrete choices?

Microeconomic theory of discrete goods

Expanding the microeconomic framework

- Continuous goods
- and discrete goods

The consumer

- selects the quantities of continuous goods: $q = (q_1, \ldots, q_L)$
- chooses an alternative in a discrete choice set $i = 1, \ldots, j, \ldots, J$
- discrete decision vector: $(w_1, \ldots, w_J)$, $w_j \in \{0, 1\}$, $\sum_j w_j = 1$.

Utility maximization

Utility

$$\widetilde{U}(q, w, \tilde{z}^T w; \theta)$$

- q: quantities of the continuous good
- w: discrete choice
- $\tilde{z}^T = (\tilde{z}_1, \ldots, \tilde{z}_i, \ldots, \tilde{z}_J) \in \mathbb{R}^{K \times J}$: K attributes of the J alternatives
- $\tilde{z}^T w \in \mathbb{R}^K$: attributes of the chosen alternative
- θ: vector of parameters

Utility maximization

Optimization problem

$$\max_{q,w} \widetilde{U}(q, w, \tilde{z}^T w; \theta)$$

subject to

$$\begin{aligned} & p^T q + c^T w \leq I \\ & \textstyle\sum_j w_j = 1 \\ & w_j \in \{0, 1\}, \forall j. \end{aligned}$$

where $c^T = (c_1, \ldots, c_i, \ldots, c_J)$ contains the cost of each alternative.

Solving the problem

- Mixed integer optimization problem
- No optimality condition
- Impossible to derive demand functions directly

Solving the problem

Step 1: condition on the choice of the discrete good

- Fix the discrete good, that is select a feasible w.
- The problem becomes a continuous problem in q.
- Conditional demand functions can be derived:

$$q_{\ell|w} = f(I - c^T w, p, \tilde{z}^T w; \theta),$$

  or, equivalently, for each alternative i,

$$q_{\ell|i} = f(I - c_i, p, \tilde{z}_i; \theta).$$

- $I - c_i$ is the income left for the continuous goods, if alternative i is chosen.
- If $I - c_i < 0$, alternative i is declared unavailable and removed from the choice set.

Solving the problem

Conditional indirect utility functions

Substitute the demand functions into the utility:

$$U_i = U(I - c_i, p, \tilde{z}_i; \theta) \text{ for all } i \in \mathcal{C}.$$

Step 2: Choice of the discrete good

$$\max_w U(I - c^T w, p, \tilde{z}^T w; \theta)$$

- Enumerate all alternatives.
- Compute the conditional indirect utility function U_i.
- Select the alternative with the highest U_i.
- Note: no income constraint anymore.

Simple example: mode choice

Attributes

	Attributes	
Alternatives	Travel time (t)	Travel cost (c)
Car (1)	t_1	c_1
Bus (2)	t_2	c_2

Utility

$$\widetilde{U} = \widetilde{U}(w_1, w_2),$$

where we impose the restrictions that, for $i = 1, 2$,

$$w_i = \begin{cases} 1 & \text{if travel alternative i is chosen,} \\ 0 & \text{otherwise;} \end{cases}$$

and that only one alternative is chosen: $w_1 + w_2 = 1$.

Simple example: mode choice

Utility functions

$$\begin{aligned} U_1 &= -\beta_t t_1 - \beta_c c_1, \\ U_2 &= -\beta_t t_2 - \beta_c c_2, \end{aligned}$$

where $\beta_t > 0$ and $\beta_c > 0$ are parameters.

Equivalent specification

$$\begin{aligned} U_1 &= -(\beta_t/\beta_c) t_1 - c_1 &= -\beta t_1 - c_1 \\ U_2 &= -(\beta_t/\beta_c) t_2 - c_2 &= -\beta t_2 - c_2 \end{aligned}$$

where $\beta > 0$ is a parameter.

Choice

- Alternative 1 is chosen if $U_1 \geq U_2$.
- Ties are ignored.

Simple example: mode choice

Choice

Alternative 1 is chosen if

$$-\beta t_1 - c_1 \geq -\beta t_2 - c_2$$

or

$$-\beta(t_1 - t_2) \geq c_1 - c_2$$

Alternative 2 is chosen if

$$-\beta t_1 - c_1 \leq -\beta t_2 - c_2$$

or

$$-\beta(t_1 - t_2) \leq c_1 - c_2$$

Dominated alternative

- If $c_2 > c_1$ and $t_2 > t_1$, $U_1 > U_2$ for any $\beta > 0$
- If $c_1 > c_2$ and $t_1 > t_2$, $U_2 > U_1$ for any $\beta > 0$

Simple example: mode choice

Trade-off

- Assume $c_2 > c_1$ and $t_1 > t_2$.
- Is the traveler willing to pay the extra cost $c_2 - c_1$ to save the extra time $t_1 - t_2$?
- Alternative 2 is chosen if

$$-\beta(t_1 - t_2) \leq c_1 - c_2$$

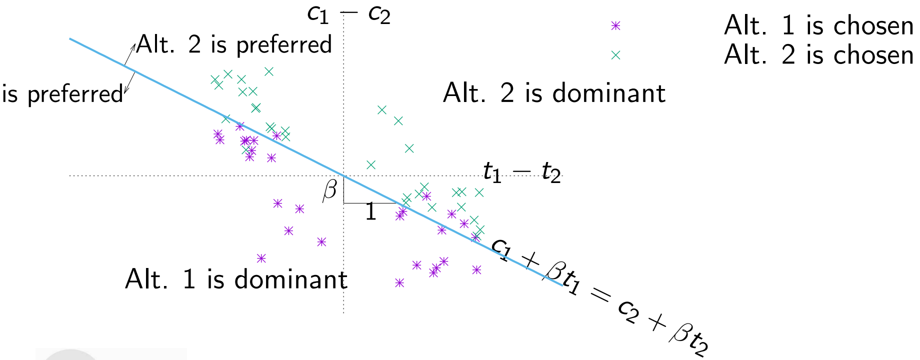
or

$$\beta \geq \frac{c_2 - c_1}{t_1 - t_2}$$

- β is called the willingness to pay or value of time

TRANSP-OR EPFL

Simple example: mode choice



Behavioral validity of the utility maximization?

Assumptions

Decision-makers

- are able to process information
- have perfect discrimination power
- have transitive preferences
- are perfect maximizer
- are always consistent

Relax the assumptions

Use a probabilistic approach: what is the probability that alternative i is chosen?

Random utility model

Probability model

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

Random utility

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

Similarity with linear regression

$$Y = \beta^T X + \varepsilon$$

Here, U is not observed. Only the choice is observed.

Derivation

Joint distributions of ε_n

- Assume that $\varepsilon_n = (\varepsilon_{1n}, \ldots, \varepsilon_{J_n n})$ is a multivariate random variable
- with CDF

$$F_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n})$$

- and pdf

$$f_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n}) = \frac{\partial^{J_n} F}{\partial \varepsilon_1 \cdots \partial \varepsilon_{J_n}}(\varepsilon_1, \ldots, \varepsilon_{J_n}).$$

The random utility model: $P_n(i|\mathcal{C}_n) =$

$$\int_{\varepsilon=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) d\varepsilon$$

Random utility model

Logit model

- The general formulation is complex.
- Assuming that ε_{in} are i.i.d. EV(0, μ), we have the logit model:

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

Outline

A simple example

Data

- $\mathcal{C}$: set of movies
- Population of N individuals
- Utility function:
  $U_{in} = \beta_{in} p_{in} + f(z_{in}) + \varepsilon_{in}$

Decision variables

- What movies to propose? y_{in}
- What price? p_{in}

Profit maximization

Data

- Two alternatives: my theater (m) and the competition (c)
- We assume an heterogenous population of N individuals

$$U_{cn} = 0 + \varepsilon_{cn}$$
$$U_{mn} = \beta_n p_m + c_{mn} + \varepsilon_{mn}$$

- $\beta_n < 0$
- Logit model: ε_{mn} i.i.d. EV

TRANSP-OR EPFL

Heterogeneous population

Two groups in the population

$$U_{mn} = \beta_n p_m + c_{mn} + \varepsilon_{mn}$$

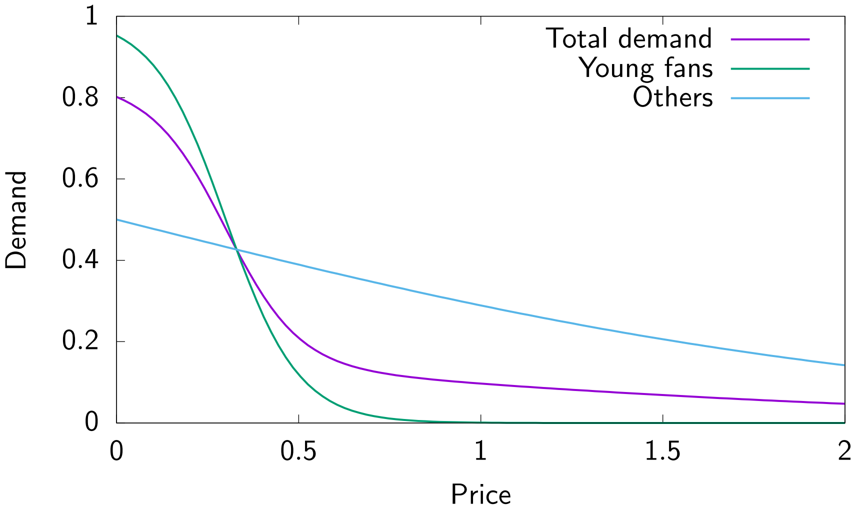
$n = 1$: Young fans: 2/3

$\beta_1 = -10$, $c_{m1} = 3$

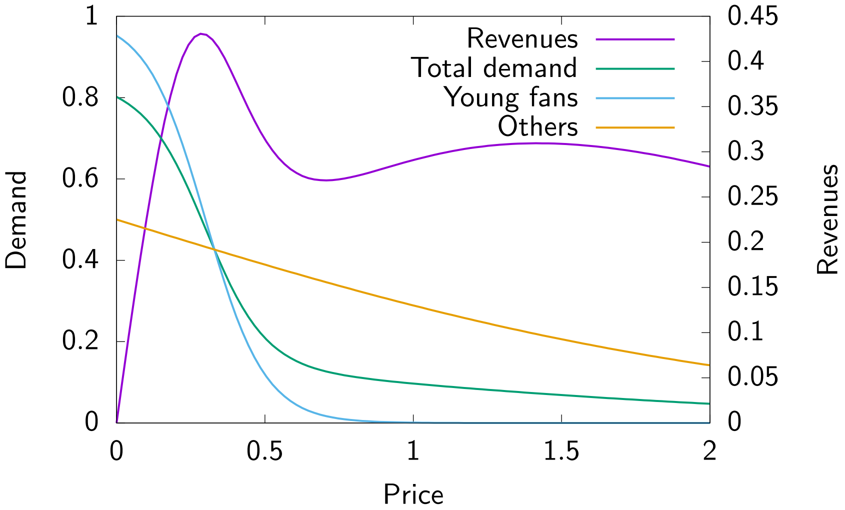
$n = 2$: Others: 1/3

$\beta_2 = -0.9$, $c_{m2} = 0$

Demand


1
0.8
0.6
0.4
0.2
0
Demand
0
0.5
1
1.5
2
Price
Total demand
Young fans
Others

Demand and revenues



Optimization

Profit maximization

- Non linear
- Non convex

Solution: mathematical programming

- Random term: simulation.
- Utility maximization of customers: constraints.

Utility

Variables

U_{inr} utility

$$z_{inr} = \begin{cases} U_{inr} & \text{if } y_{in} = 1 \\ \ell_{nr} & \text{if } y_{in} = 0 \end{cases}$$ discounted utility

(ℓ_{nr} smallest lower bound)

Constraint: utility

$$U_{inr} = \overbrace{\beta_{in} p_{in} + q_d(x_d)}^{V_{in}} + \xi_{inr} \; \forall i, n, r$$

Utility (ctd)

Constraints: discounted utility

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

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

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

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

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

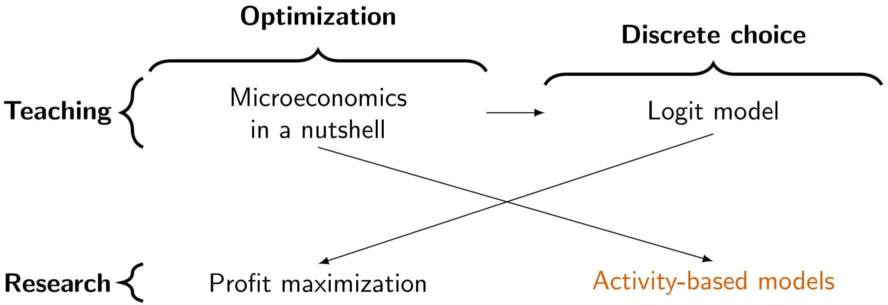
$$w_{inr} \leq y_{in} \qquad \forall i, n, r$$

Profit maximization problem

MILP

- We avoid the non convex formulation of the logit model.
- Most constraints are linear.
- Nonlinear constraints are easy to linearize.
- No specific assumption of the distribution of ε_{in} thanks to simulation.
- Very large optimization problems.
- Current research: decomposition methods (Benders, column-generation, etc.)

Summary


Optimization
Discrete choice
Teaching
Microeconomics in a nutshell
Logit model
Research
Profit maximization
Activity-based models

Bibliography I

- Ben-Akiva, M. and Bierlaire, M. (2003).
  Discrete choice models with applications to departure time and route choice.
  Operations Research and Management Science, pages 7–38. Kluwer.

- Bortolomiol, S. (2022).
  Optimization and equilibrium problems with discrete choice models of demand.
  PhD thesis, EPFL.

- Bortolomiol, S., Lurkin, V., and Bierlaire, M. (2021a).
  Price-based regulation of oligopolistic markets under discrete choice models of demand.
  Transportation.
  Accepted on Jul 23, 2021.

Bibliography II

Bortolomiol, S., Lurkin, V., and Bierlaire, M. (2021b).
A simulation-based heuristic to find approximate equilibria with disaggregate demand models.
Transportation Science, 55(5):1025–1045.
Accepted on Apr 16, 2021.

Pacheco, M. (2020).
A general framework for the integration of complex choice models into mixed integer optimization.
PhD thesis, EPFL.

Bibliography III

Pacheco, M., Bierlaire, M., Gendron, B., and Sharif Azadeh, S. (2021).
Integrating advanced discrete choice models in mixed integer linear optimization.
Transportation Research Part B: Methodological, 146:26–49.
Accepted on Feb 09, 2021.