Discrete choice models and operations research: a difficult combination

Michel Bierlaire  Shadi Sharif Azadeh

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

June 16, 2016
Outline

1. Demand and supply
2. Disaggregate demand models
3. Choice-based optimization
   - Applications
4. A generic framework
5. A simple example
   - Example: one theater
   - Example: two theaters
   - Example: two theaters with capacities
6. Conclusion
Demand and supply

**Demand models**

- Supply = infrastructure
- Demand = behavior, choices
- Congestion = mismatch
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: $Q = f^{-1}(P)$
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.
Demand-supply interactions

Operations Research
- Given the demand...
- configure the system

Behavioral models
- Given the configuration of the system...
- predict the demand
Demand-supply interactions

Multi-objective optimization

Minimize costs

Maximize satisfaction
Outline

1. Demand and supply
2. Disaggregate demand models
3. Choice-based optimization
   - Applications
4. A generic framework
5. A simple example
   - Example: one theater
   - Example: two theaters
   - Example: two theaters with capacities
6. Conclusion
Choice models

**Behavioral models**
- Demand = sequence of choices
- Choosing means trade-offs
- In practice: derive trade-offs from choice models
Theoretical foundations

- Random utility theory
- Choice set: $C_n$
- $y_{in} = 1$ if $i \in C_n$, 0 if not
- Logit model:

$$P(i|C_n) = \frac{y_{in}e^{V_{in}}}{\sum_{j \in C} y_{jn}e^{V_{jn}}}$$
Logit model

Utility

\[ U_{in} = V_{in} + \varepsilon_{in} \]

- Decision-maker \( n \)
- Alternative \( i \in C_n \)

Choice probability

\[ P_n(i|C_n) = \frac{y_{in}e^{V_{in}}}{\sum_{j \in C} y_{jn}e^{V_{jn}}} \]
Variables: $x_{in} = (z_{in}, s_n)$

Attributes of alternative $i$: $z_{in}$
- Cost / price
- Travel time
- Waiting time
- Level of comfort
- Number of transfers
- Late/early arrival
- etc.

Characteristics of decision-maker $n$: $s_n$
- Income
- Age
- Sex
- Trip purpose
- Car ownership
- Education
- Profession
- etc.
Demand curve

Disaggregate model

\[ P_n(i|c_{in}, z_{in}, s_n) \]

Total demand

\[ D(i) = \sum_n P_n(i|c_{in}, z_{in}, s_n) \]

Difficulty

Non linear and non convex in \( c_{in} \) and \( z_{in} \)
Outline

1. Demand and supply
2. Disaggregate demand models
3. Choice-based optimization
   - Applications
4. A generic framework
5. A simple example
   - Example: one theater
   - Example: two theaters
   - Example: two theaters with capacities
6. Conclusion
Choice-Based Optimization Models

Benefits
- Merging supply and demand aspect of planning
- Accounting for the heterogeneity of demand
- Dealing with complex substitution patterns
- Investigation of demand elasticity against its main driver (e.g. price)

Challenges
- Nonlinearity and nonconvexity
- Assumptions for simple models (logit) may be inappropriate
- Advanced demand models have no closed-form
- Endogeneity: same variable(s) both in the demand function and the cost function
Stochastic traffic assignment

Features
- Nash equilibrium
- Flow problem
- Demand: path choice
- Supply: capacity
Selected literature

- [Dial, 1971]: logit
- [Daganzo and Sheffi, 1977]: probit
- [Fisk, 1980]: logit
- [Bekhor and Prashker, 2001]: cross-nested logit
- and many others...
Revenue management

Features

- Stackelberg game
- Bi-level optimization
- Demand: purchase
- Supply: price and capacity
Selected literature

- [Labbé et al., 1998]: bi-level programming
- [Andersson, 1998]: choice-based RM
- [Talluri and Van Ryzin, 2004]: choice-based RM
- [Gilbert et al., 2014a]: logit
- [Gilbert et al., 2014b]: mixed logit
- [Azadeh et al., 2015]: global optimization
- and many others...
Facility location problem

Features

- Competitive market
- Opening a facility impact the costs
- Opening a facility impact the demand
- Decision variables: availability of the alternatives

\[ P_n(i|C_n) = \frac{y_{in}e^{V_{in}}}{\sum_{j \in C} y_{jn}e^{V_{jn}}}. \]
Selected literature

- [Hakimi, 1990]: competitive location (heuristics)
- [Benati, 1999]: competitive location (B & B, Lagrangian relaxation, submodularity)
- [Serra and Colomé, 2001]: competitive location (heuristics)
- [Marianov et al., 2008]: competitive location (heuristic)
- [Haase and Müller, 2013]: school location (simulation-based)
The main idea... during my sabbatical in Montréal
The main idea

Linearization
Hopeless to linearize the logit formula (we tried...)

First principles
Each customer solves an optimization problem

Solution
Use the utility and not the probability
A linear formulation

Utility function

\[ U_{in} = V_{in} + \varepsilon_{in} = \sum_{k} \beta_k x_{ink} + f(z_{in}) + \varepsilon_{in}. \]

Simulation

- Assume a distribution for \( \varepsilon_{in} \)
- E.g. logit: i.i.d. extreme value
- Draw \( R \) realizations \( \xi_{inr}, \) \( r = 1, \ldots, R \)
- The choice problem becomes deterministic
**Scenarios**

**Draws**

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

$$U_{inr} = \sum_k \beta_k x_{ink} + f(z_{in}) + \xi_{inr}.$$  

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

Variables

Availability

\[ y_{in} = \begin{cases} 1 & \text{if alt. } i \text{ available for } n, \\ 0 & \text{otherwise.} \end{cases} \]

Choice

\[ w_{inr} = \begin{cases} 1 & \text{if } y_{in} = 1 \text{ and } U_{inr} = \max_{j \mid y_{jn} = 1} U_{jn}, \\ 0 & \text{if } y_{in} = 0 \text{ or } U_{inr} < \max_{j \mid y_{jn} = 1} U_{jn}. \end{cases} \]
A generic framework

## Capacities

- Demand may exceed supply
- Each alternative $i$ can be chosen by maximum $c_i$ individuals.
- An exogenous priority list is available.
- The numbering of individuals is consistent with their priority.
Priority list

Application dependent
- First in, first out
- Frequent travelers
- Subscribers
- ...

In this framework
The list of customers must be sorted
Capacities

Variables

- $y_{in}$: decision of the operator
- $y_{inr}$: availability

Constraints

\[
\sum_{i \in C} w_{inr} = 1 \quad \forall n, r.
\]

\[
\sum_{n=1}^{N} w_{inr} \leq c_i \quad \forall i, n, r.
\]

\[
w_{inr} \leq y_{inr} \quad \forall i, n, r.
\]

\[
y_{inr} \leq y_{in} \quad \forall i, n, r.
\]

\[
y_{i(n+1)r} \leq y_{inr} \quad \forall i, n, r.
\]
Demand and revenues

Demand

\[ D_i = \frac{1}{R} \sum_{n=1}^{N} \sum_{r=1}^{R} w_{inr}. \]

Revenues

\[ R_i = \frac{1}{R} \sum_{n=1}^{N} p_{in} \sum_{r=1}^{R} w_{inr}. \]
Revenues

Non linear specification

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

Linearization

Predetermined price levels

Price levels: \( p_{in}^{\ell}, \ell = 1, \ldots, L_{in} \)

\[ p_{in} = \sum_{\ell=1}^{L_{in}} \lambda_{in\ell} p_{in}^{\ell}. \]

New decision variables

\( \lambda_{in\ell} \in \{0, 1\} \)

\[ \sum_{\ell=1}^{L_{in}} \lambda_{in\ell} = 1. \]
References

- Technical report: [Bierlaire and Azadeh, 2016]
- Conference proceeding: [Pacheco et al., 2016a]
- TRISTAN presentation: [Pacheco et al., 2016b]
A simple example

Data
- $C$: 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? $y_i$
- What price? $p_{in}$
Back to the example: pricing

Data

- Two alternatives: my theater \((m)\) and the competition \((c)\)

- We assume an homogeneous population of \(N\) individuals

\[
U_c = 0 + \varepsilon_c \\
U_m = \beta_c p_m + \varepsilon_m
\]

- \(\beta_c < 0\)

- Logit model: \(\varepsilon_m\) i.i.d. EV
Demand and revenues
Optimization (with GLPK)

Data

- $N = 1$
- $R = 100$
- $U_m = -10p_m + 3$
- Prices: 0.10, 0.20, 0.30, 0.40, 0.50

Results

- Optimum price: 0.3
- Demand: 56%
- Revenues: 0.168
Heterogeneous population

Two groups in the population

\[ U_{in} = -\beta_n p_i + c_n \]

Young fans: 2/3
\[ \beta_1 = -10, \quad c_1 = 3 \]

Others: 1/3
\[ \beta_1 = -0.9, \quad c_1 = 0 \]
Demand and revenues
A simple example
Example: one theater

Optimization

Data

- $N = 3$
- $R = 100$
- $U_{m1} = -10p_m + 3$
- $U_{m2} = -0.9p_m$
- Prices: 0.3, 0.7, 1.1, 1.5, 1.9

Results

- Optimum price: 0.3
- Customer 1 (fan): 60% [theory: 50%]
- Customer 2 (fan): 49% [theory: 50%]
- Customer 3 (other): 45% [theory: 43%]
- Demand: 1.54 (51%)
- Revenues: 0.48
Two theaters, different types of films
Two theaters, different types of films

**Theater $m$**
- Expensive
- Star Wars Episode VII

**Theater $k$**
- Cheap
- Tinker Tailor Soldier Spy

**Heterogeneous demand**
- Two third of the population is young (price sensitive)
- One third of the population is old (less price sensitive)
Two theaters, different types of films

Data

- Theaters $m$ and $k$
- $N = 6$
- $R = 10$
- $U_{mn} = -10p_m + 4$, $n = 1, 2, 4, 5$
- $U_{mn} = -0.9p_m$, $n = 3, 6$
- $U_{kn} = -10p_k + 0$, $n = 1, 2, 4, 5$
- $U_{kn} = -0.9p_k$, $n = 3, 6$
- Prices $m$: 1.0, 1.2, 1.4, 1.6, 1.8
- Prices $k$: half price

Theater $m$

- Optimum price $m$: 1.6
- 4 young customers: 0
- 2 old customers: 0.5
- Demand: 0.5 (8.3%)
- Revenues: 0.8

Theater $k$

- Optimum price $m$: 0.5
- Young customers: 0.8
- Old customers: 1.5
- Demand: 2.3 (38%)
- Revenues: 1.15
Two theaters, same type of films

**Theater $m$**
- Expensive
- Star Wars Episode VII

**Theater $k$**
- Cheap
- Star Wars Episode VIII

**Heterogeneous demand**
- Two third of the population is young (price sensitive)
- One third of the population is old (less price sensitive)
Two theaters, same type of films

Data

- Theaters $m$ and $k$
- $N = 6$
- $R = 10$
- $U_{mn} = -10p_m + 4$, $n = 1, 2, 4, 5$
- $U_{mn} = -0.9p_m$, $n = 3, 6$
- $U_{kn} = -10p_k + 4$, $n = 1, 2, 4, 5$
- $U_{kn} = -0.9p_k$, $n = 3, 6$
- Prices $m$: 1.0, 1.2, 1.4, 1.6, 1.8
- Prices $k$: half price

Theater $m$

- Optimum price $m$: 1.8
- Young customers: 0
- Old customers: 1.9
- Demand: 1.9 (31.7%)
- Revenues: 3.42

Theater $k$

Closed
Two theaters with capacity, different types of films

Data
- Theaters $m$ and $k$
- Capacity: 2
- $N = 6$
- $R = 5$
- $U_{mn} = -10p_m + 4, \ n = 1, 2, 4, 5$
- $U_{mn} = -0.9p_m, \ n = 3, 6$
- $U_{kn} = -10p_k + 0, \ n = 1, 2, 4, 5$
- $U_{kn} = -0.9p_k, \ n = 3, 6$
- Prices $m$: 1.0, 1.2, 1.4, 1.6, 1.8
- Prices $k$: half price

Theater $m$
- Optimum price $m$: 1.8
- Demand: 0.2 (3.3%)
- Revenues: 0.36

Theater $k$
- Optimum price $m$: 0.5
- Demand: 2 (33.3%)
- Revenues: 1.15
### Example of two scenarios

<table>
<thead>
<tr>
<th>Customer</th>
<th>Choice</th>
<th>Capacity $m$</th>
<th>Capacity $k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>$k$</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>$k$</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customer</th>
<th>Choice</th>
<th>Capacity $m$</th>
<th>Capacity $k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>$k$</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>$k$</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
Outline

1. Demand and supply
2. Disaggregate demand models
3. Choice-based optimization
   - Applications
4. A generic framework
5. A simple example
   - Example: one theater
   - Example: two theaters
   - Example: two theaters with capacities
6. Conclusion
Summary

Demand and supply
- Supply: prices and capacity
- Demand: choice of customers
- Interaction between the two

Discrete choice models
- Rich family of behavioral models
- Strong theoretical foundations
- Great deal of concrete applications
- Capture the heterogeneity of behavior
- Probabilistic models
Optimization

Discrete choice models
- Non linear and non convex
- Idea: use utility instead of probability
- Rely on simulation to capture stochasticity

Proposed formulation
- Linear in the decision variables
- Large scale
- Fairly general
Ongoing research

- Decomposition methods
- Scenarios are (almost) independent from each other (except objective function)
- Individuals are also loosely coupled (except for capacity constraints)
Thank you!


Bibliography III


Bibliography VI
