Demand Estimation under Uncertain Consideration Sets

Srikanth Jagabathula
Associate Professor of Tech, Ops, & Stats
NYU Stern School of Business, New York
@s_jagabathula
http://srikanth-jagabathula.com

NYC Ops Day, Cornell Tech 2022
Demand Estimation Under Uncertain Consideration Sets

Gustavo Vulcano
Professor of OM
School of Business
Torquato di Tella University
Argentina

Dmitry Mitrofanov
Asst. Prof. in Boston College
(former PhD student @ Stern)
Discrete choice models predict demand using both what *was* and *was not* chosen.

Critical assumption in choice models:
- Purchased \( \succ \) Not purchased but available

Consideration set

Offer set
E.g., spatial choice in sharing economy platforms

available listings determined by the MARKET
Concrete application: demand as a function of prices at a P2P short-term car rental platform

- Rental prices
- Consideration set
- Rented car
- Customer location

image: A screenshot of a P2P car rental platform with search filters for location, start and end time, duration, duration, price, make, category, class, transmission, and additional options like all makes, all categories, dedicated parking spot. The search results show two rental options: a 2013 Mercedes-Benz GLK-Class for $107.45 and a 2018 Mercedes-Benz C-Class for $120.15.
Discrete choice models predict demand using both what *was* and *was not* chosen.

Consideration set ≠ offer set

UNOBSERVED consideration sets

Critical assumption in choice models:
- Purchased > NOT purchased but available

Consider-then-Choose (CTC) models:
- Sample a consideration set
- Choose from sampled set using a choice model

Standard approach: consideration set = offer set
When (if at all) should we fit CTC models over standard choice models when consideration sets are latent?

RESEARCH QUESTIONS:
When consideration sets are unobserved,
Are CTC models identified?
Are CTC models better than standard approach?

OUR CONTRIBUTIONS:
Identification conditions for a general class of CTC models
Methodology to estimate CTC models from transactions
CTC models better when future offer sets are uncertain
- long-term forecasts in retail [random stockouts]
- demand forecasts in online platforms [mkt. determined offer sets]
Related work: Rich work spanning various areas

### Marketing
[Mahajan, van Campbell ‘69], [Howard and Sheth ‘69], [Wright and Barbour ‘77], [Swait and Ben-Akiva ‘87], [Lynch ‘91], [Roberts and Lattin ‘97], [Gilbride et al. ‘04], [Hauser ‘14], ...

### Economics
[Manzini and Mariotti ‘14], [Masatlioglu et al. ‘16], [Eliaz and Spiegler ‘11], [Caplin et al. ‘16], [Matejka and McKay ‘15], [Cooper ‘88], [Blattberg and Wisniewski ‘89], ...

### Operations
[Aouad et al. ‘15], [Feldman and Topaloglu ‘15], [Wang and Sahin ‘17], [Feldman et al. ‘18], [Aouad et al. ‘19], [Li and Netessine ‘17], [Jagabathula, Rusmevichientong ‘17], [Jagabathula and Vulcano ‘18], [Lederman et al. ’14], [Gallego and Li ‘17]...

- introduce notion of consideration set
- analyze consideration set heuristics

We study product-space models, which scale to large-scale transaction data

- specialized models + theoretical results
- not geared towards predictions

We derive similar theoretical results + demonstrate predictive accuracy

- assortment + price optimization
- estimation of price-based consid. sets

We fit more general models + study noise in availability sets
Sales transaction data: choice + offer set

$n$ products

mutually substitutable

Observations: for each purchase instance $t$

<table>
<thead>
<tr>
<th>offer set $S_t$ = filled dots</th>
<th>chosen product $i_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 1$</td>
<td></td>
</tr>
<tr>
<td>$S_1 = {1,2,3,4,6,7}$, $i_1 = 3$</td>
<td></td>
</tr>
<tr>
<td>$t = 2$</td>
<td></td>
</tr>
<tr>
<td>$S_2 = {1,2,3,6,7}$, $i_2 = 3$</td>
<td></td>
</tr>
<tr>
<td>$t = 3$</td>
<td></td>
</tr>
<tr>
<td>$S_3 = {1,2,6,7}$, $i_3 = 2$</td>
<td></td>
</tr>
</tbody>
</table>
Choice model: customers sample consideration sets and choose from them

Customers characterized by:
- distribution over product preference lists
- distribution \( \lambda \) over subsets of products

In each choice instance, customers:
- samples a preference list
- samples a consideration set
- purchases most preferred considered item that is offered

Consideration set:
- samples from distribution \( \lambda \)
- decreasing order of preference

Offer set:
- purchased item
- items not considered
- items considered but not offered
**General Consider-then-Choose (GCC) choice model**

Customers characterized by a single product preference list:
- Distribution $\lambda$ over subsets.

In each choice instance:
- Samples a preference list.
- Samples a consideration set.
- Purchases most preferred considered item that is offered.

Consideration set sampled from distribution $\lambda$ in decreasing order of preference.

Offer set: 2, 3, 4, 5, 6, 7.

1. Considered but not offered: 1, 2, 5.
2. Not considered: 3.
Choice model: We consider product-space consideration set models

Models in product space:

Independent model: product $j$ considered w/ prob. $\theta_j$

\[
\lambda(C) = \prod_{j \in C} \theta_j \prod_{j \notin C} (1 - \theta_j)
\]

[Manzini and Mariotti ’14] [Gallego and Li ‘17]

General model: $\lambda$ is a general dist. over subsets

\[
\lambda(C) = \sum_{k=1}^{K} \prod_{j \in C} \theta_{jk} \prod_{j \notin C} (1 - \theta_{jk})
\]

mixtures can approx. general distributions

Models in feature space:

product considered if all its features above acceptable “thresholds”

common in marketing BUT, computationally difficult

[Gilbride and Allenby ‘04]

We focus on product-space models and let $\theta_{jk}$ depend on product features

vary across cust.
GCC model vs. standard choice models?

Theorem (informal):

GCC model is a special case of the RUM choice rule. That is, all choice probabilities under a GCC model can be matched with those under a distribution over rankings assuming that consideration set = offer set.

Does it ever help to fit GCC model and not the RUM model?

Is the literature on assortment and price opt with CTC models irrelevant?
Identifying model parameters from sales transactions data

If all customers share the same preference list \( \sigma \), then model parameters are identified from sales transaction data alone.

Theorem (informal):

If all customers share the same preference list \( \sigma \), then model parameters are identified from sales transaction data alone.

\[
\lambda(C) = \sum_{X \subseteq C} (-1)^{|C| - |X|} \mathbb{P}_0(N \setminus X)
\]

product \( i \) > product \( j \) under ranking \( \sigma \) \( \iff \) \( j \) DOES NOT cannibalize sales of \( i \)
When consideration sets are “small,” model parameters can be identified using data from “small” offer sets.

If all customers share the same preference list $\sigma$, and consideration sets are of size at most $k$, then model parameters are identified from sales transaction data alone.

**Theorem (informal):**

If all customers share the same preference list $\sigma$, and consideration sets are of size at most $k$, then model parameters are identified from sales transaction data alone.

\[
\lambda(C') = \sum_{X \subseteq N} \sum_{Y \supseteq X \cup C} (-1)^{1+|Y|-|X \Delta C|} I[|X \cup C| \leq k \leq |Y|] P_0(X)
\]

product $i >$ product $j$ under ranking $\sigma \iff j$ DOES NOT cannibalize sales of $i$
Identification problem: Characterize our model through revealed preferences

**Condition 1:** One-directional cannibalization
- product $i$ cannibalizes product $j$ OR
- product $j$ cannibalizes product $i$ but NOT BOTH

**Condition 2:** Cannibalization is menu independent
If product $i$ cannibalizes product $j$ in an offer set, then it cannibalizes in all offer sets

**Condition 3:** Technical condition to ensure validity of consideration set distribution

$$\sum_{X \subseteq C} (-1)^{|C|-|X|} \mathbb{P}_0(N \setminus X) \geq 0$$
for all subsets $C$, with strict ineq. when $|C| < 4$

**Theorem (informal):**
Choice data consistent with our model iff conditions 1, 2, and 3 are satisfied
IRI Academic Data Set

Weekly store sales data

31 product categories
11 years 2001 - 2011

We analyzed year 2007 data

22 product categories
400K transactions across 52 weeks
### Experiments conducted

#### Training data
- weekly + store-level product availability data
- weekly purchases

#### Test data: three scenarios

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Prediction task</th>
</tr>
</thead>
<tbody>
<tr>
<td>short-term</td>
<td>weekly + store-level product availability data</td>
<td>market shares?</td>
</tr>
<tr>
<td>predictions</td>
<td>(aggregated across weeks in test data)</td>
<td></td>
</tr>
<tr>
<td>long-term</td>
<td>store-level product availability data</td>
<td>market shares?</td>
</tr>
<tr>
<td>predictions</td>
<td>(aggregated across weeks in test data)</td>
<td></td>
</tr>
<tr>
<td>DC-level</td>
<td>weekly product availability data</td>
<td>market shares?</td>
</tr>
<tr>
<td>predictions</td>
<td>(aggregated across stores in test data)</td>
<td></td>
</tr>
</tbody>
</table>
Results: we obtain significantly more accurate predictions for long-term and DC-level predictions.
Results: we obtain significantly more accurate predictions for long-term and DC-level predictions.
Results: improvement from our method greatest when estimated offer set differs the most from true offer set.
Online platform: Peer-to-peer car sharing service

Short-term car rental data: July 2012 – Aug 2014

- 34 car brands
- 514 car owners
- 7.5K renters
- 26.8K transactions

Car sharing cities:
- San Francisco
- Chicago
- Portland
- Other cities in California

Car features (available to the renter):
- **Car brand** (e.g., BMW, Tesla)
- **Car location type and accessibility** (e.g., car access, car location hours, car location type)
- **Car type** (e.g., Economy, SUV)
- **Other car features** (e.g., hourly price, car age, transmission, premium wheels, power seats, Bluetooth/wireless, leather interior, sunroof/moonroof, premium sound, power windows, GPS, roof rack, tinted windows)
Experiments conducted

Training data (~1.5 years)

snapshot of cars available in SF at each time

location + time of car rental

rented car
Experiments conducted

Training data (~1.5 years)
snapshot of cars available in SF at each time location + time of car rental

Prediction task  market shares of car brands over test period (~4.5 months)
(e.g., what fraction of rentals are Toyota, BMW, etc.)

Challenge  car availability determined by the market
location + times of rental unknown
Experiments conducted

GIVEN
platform growth trajectory: total # of transactions in each cell
all cars listed in the cell over the last ~2 months

PREDICT
market shares for each car make

7 x 7 grid
Results: we obtain significantly more accurate predictions for car-make market shares.

- **17%** MAPE improvement for our base method.
- **44%** MAPE improvement for our random-forest based method.

Can incorporate cutting-edge ML methods.
Results: we obtain significantly more accurate predictions for car-make market shares

% MAPE [lower is better]

% RMSE [lower is better]
Results: our method is consistently better across all makes of cars
Results: the decision tree model is intuitive and interpretable
MAIN TAKEAWAYS

Consideration sets + availability sets: often unobserved, particularly in settings in which availability determined by market.

Asymmetry between training and test sets: often better info on availability sets in training than test.

Explicit modeling of consideration sets: results in better predictions; particularly, long-term predictions.
Demand Estimation under Uncertain Consideration Sets

Srikanth Jagabathula
Associate Professor of Tech, Ops, & Stats
NYU Stern School of Business, New York

Web:  http://srikanth-jagabathula.com
Twitter:  @s_jagabathula
LinkedIn:  https://www.linkedin.com/in/srikanth-jagabathula-286a19a