The Limit of the Marginal Distribution Model in Consumer Choice

Joint work with: Xiaobo Li Karthyek Murthy Karthik Natarajan

Presented by Ruan Yanqiu

August 4, 2022

Problem Description

- Given data on choices made by consumers for different assortments, a key challenge is to develop parsimonious models that describe and predict consumer choice behavior.
- Main focus in our work
  1. Can the choice probability data under a set of different assortments represented by a choice model?
  2. What's the computational technique?
  3. If not, how can we find the best fit of a choice model to the given choice data?
  4. Given the choice probability being feasible for the model, can we predict the choice probability of an alternative in an unseen assortment?

Marginal Distribution Model

Let $\mathcal{N} = \{1, 2, ..., n\}$ be a set of products. The utility of each alternative $i \in \mathcal{N}$ takes the form $\tilde{u}_i = v_i + \tilde{\epsilon}_i$, where $\boldsymbol{v} = (v_1, ..., v_n)$ and $\tilde{\boldsymbol{\epsilon}} = (\tilde{\epsilon}_1, ..., \tilde{\epsilon}_n)$ denote the deterministic and stochastic parts of the utilities. In the marginal distribution model (MDM), the joint distribution of the random vector $\tilde{\boldsymbol{\epsilon}}$ is not specified, rather only the marginal distributions are specified. Let Θ denote this set of joint distributions θ for $\tilde{\boldsymbol{\epsilon}}$. Given an assortment S, the MDM computes the maximum expected consumer utility over all distributions in the set:

$$\sup_{\theta \in \Theta} \mathbb{E}_{\tilde{\boldsymbol{\epsilon}} \sim \theta} \left[\max_{i \in S} v_i + \tilde{\epsilon}_i \right]. \tag{1}$$

The choice probability $x_i^* = \mathbb{P}_{\tilde{\boldsymbol{\epsilon}} \sim \theta^*}(i = \arg\max_{j \in S} v_j + \tilde{\epsilon}_j)$ is evaluated for the distribution θ^* which attains the maximum in (1). Under appropriate assumptions, the convex formulation of computing the choice probabilities for a distribution which attains the maximum in (1)

$$\max \left\{ \sum_{i \in S} v_i x_i + \sum_{i \in S} \int_{1 - x_i}^{1} F_i^{-1}(t)\, dt \;\middle|\; \sum_{i \in S} x_i = 1, x_i \geq 0 \;\; \forall i \in S \right\}, \tag{2}$$

with the convention that $F_i^{-1}(0) = \lim_{t \downarrow 0} F_i^{-1}(t)$ and $F_i^{-1}(1) = \lim_{t \uparrow 1} F_i^{-1}(t)$.

Group Marginal Distribution Model (G-MDM)

Assumption 1

There exists a partition $\mathscr{G} = \{G_1, G_2, \ldots, G_K\}$ of the set of alternatives $\mathscr{N}$ such that the marginal distribution of $\tilde{\epsilon}_i$, for any $i \in G_l$ is given by $F_l(\cdot)$.

Let $g : \mathscr{N} \mapsto \mathscr{G}$ be a function that maps an alternative in $\mathscr{N}$ to a group in $\mathscr{G}$. For example, $g(i) = l$ means $i \in G_l$. Then one can write the objective function in the convex formulation for G-MDM as:

$$\sum_{i \in S} \nu_i x_i + \sum_{i \in S} \int_{1-x_i}^{1} F_{g(i)}^{-1}(t)\, dt = \sum_{i \in S} \nu_i x_i + \sum_{l=1}^{K} \sum_{i \in S : g(i) = l} \int_{1-x_i}^{1} F_l^{-1}(t)\, dt. \tag{3}$$

Main Theorems

Theorem 1(Feasibility conditions for G-MDM) A choice probability collection $\boldsymbol{p}_{\mathscr{S}}$ is feasible for G-MDM satisfying Assumptions 1 - 2 if and only if there exists $\boldsymbol{\lambda} \in \mathbb{R}^m, \boldsymbol{v} \in \mathbb{R}^n$ such that for all $(i, S), (j, T) \in \mathscr{I}$ with $g(i) = g(j)$:

$$\lambda_S - v_i > \lambda_T - v_j \text{ if } p_{i,S} < p_{j,T} \text{ and } \lambda_S - v_i = \lambda_T - v_j \text{ if } 0 < p_{i,S} = p_{j,T}.$$

Checking these conditions is possible in polynomial time.

Theorem 2(Representation power of MDM and RUM) When $n = 2$, RUM and MDM are equivalent; when $n = 3$, RUM subsumes MDM; when $n \geq 4$, there exist choice probabilities that can be represented by both RUM and MDM and neither of the models subsumes the other.

Theorem 3(Choice probability prediction interval) Given a G-MDM instance $(\boldsymbol{x}_{\mathscr{S}}, \boldsymbol{\lambda}, \boldsymbol{v})$ and a product k from a new assortment $A \notin \mathscr{S}$, a prediction interval for the choice probability $x_{k,A}$ which includes all G-MDM consistent with the instance $(\boldsymbol{x}_{\mathscr{S}}, \boldsymbol{\lambda}, \boldsymbol{v})$ is given by,

$$\left[1 - \bar{F}_{g(k)}(\bar{\lambda} - v_k),\ 1 - \underline{F}_{g(k)}(\underline{\lambda} - v_k)\right]$$

where $\underline{\lambda}$ and $\bar{\lambda}$, respectively, are the supremum and infimum values of the collection of $\lambda \in \mathbb{R}$ satisfying

$$\sum_{i \in A} \left[1 - \bar{F}_{g(i)}(\lambda - v_i)\right] \leq 1 \leq \sum_{i \in A} \left[1 - \underline{F}_{g(i)}(\lambda - v_i)\right].$$

Limit of G-MDM: Fitting the best model to choice data

We define the limit of the G-MDM as the smallest value of loss$(\boldsymbol{p}_{\mathscr{S}}, \boldsymbol{x}_{\mathscr{S}})$ attainable by fitting observed data with G-MDM. The formulation of computing the limit of the G-MDM can be formulated as a mixed integer convex program:

$$
\begin{aligned}
\min_{\boldsymbol{x}_{\mathscr{S}}, \boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\delta}, \boldsymbol{y}} \quad & \sum_{S \in \mathscr{S}} \text{loss}(\boldsymbol{p}_S, \boldsymbol{x}_S) \\
s.t. \quad & -\delta_{i,j,S,T} \leq x_{i,S} - x_{j,T} \leq 1 - (1+\epsilon)\delta_{i,j,S,T}, \quad \forall (i,S),(j,T) \in \mathscr{I} : g(i) = g(j), \\
& \lambda_S - \nu_i - \lambda_T + \nu_j \geq -1 + (1+\epsilon)\delta_{i,j,S,T}, \quad \forall (i,S),(j,T) \in \mathscr{I} : g(i) = g(j), \\
& x_{i,S} \leq 1 - y_{i,S}, \quad \forall (i,S) \in \mathscr{I}, \\
& -y_{i,S} - y_{j,T} - \delta_{i,j,S,T} - \delta_{j,i,T,S} \leq \lambda_S - \nu_i - \lambda_T + \nu_j, \quad \forall (i,S),(j,T) \in \mathscr{I} : g(i) = g(j) \\
& \lambda_S - \nu_i - \lambda_T + \nu_j \leq y_{i,S} + y_{j,T} + \delta_{i,j,S,T} + \delta_{j,i,T,S}, \quad \forall (i,S),(j,T) \in \mathscr{I} : g(i) = g(j), \\
& \sum_{i \in S} x_{i,S} = 1, \quad \forall S \in \mathscr{S}, \\
& x_{i,S} \geq 0, \quad \forall (i,S) \in \mathscr{I}, \\
& 0 \leq \lambda_S - \nu_i \leq 1, \quad \forall (i,S) \in \mathscr{I}, \\
& \delta_{i,j,S,T} \in \{0,1\}, \quad \forall (i,S),(j,T) \in \mathscr{I} : g(i) = g(j), \\
& y_{i,S} \in \{0,1\}, \quad \forall (i,S) \in \mathscr{I},
\end{aligned}
$$

for some small positive number $\epsilon > 0$.

Numerical Experiments

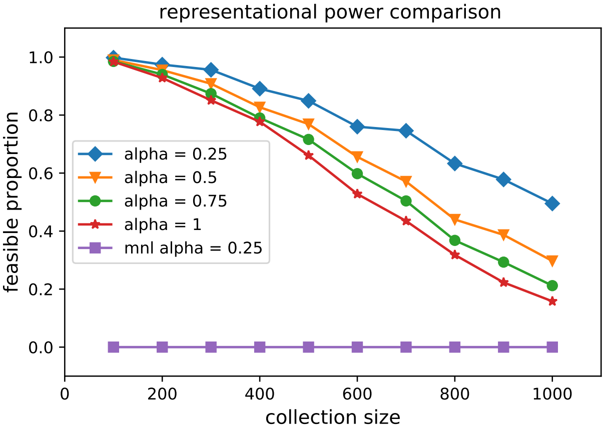

Figure: The representational power of MDM in Experiment 1

Numerical Experiments

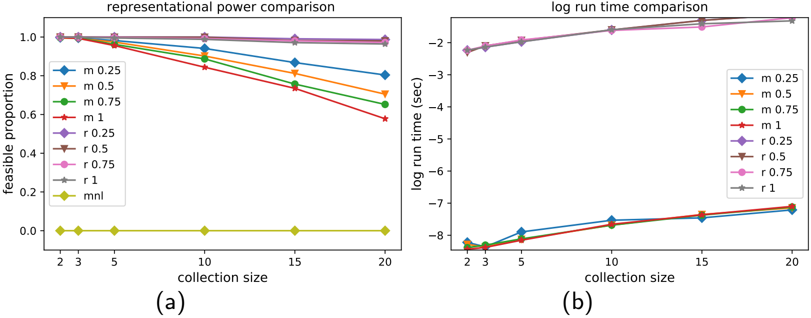

Figure: Comparison of the performance of MDM and RUM in Experiment 2, m stands for MDM, r stands for RUM, numbers stand for the perturbation parameters

Numerical Experiments

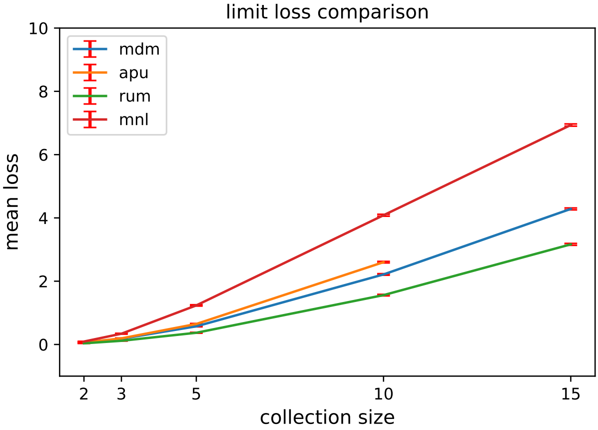

Figure: The explanatory ability comparison

Numerical Experiments

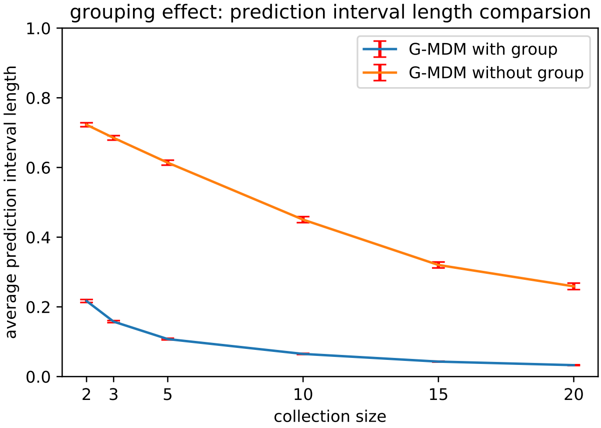

Figure: The grouping effect in G-MDM

Conclusions

- We propose a group marginal distribution model (G-MDM). Given choice data over a collection of assortments, we develop necessary and sufficient conditions to verify if a G-MDM can represent the observed choice probabilities. Unlike RUM, checking these conditions is possible in polynomial time.
- By computing the limit problem, one can obtain an MDM which offers the best fit to given choice data. This estimation approach is novel, contrasting with existing approaches which need to make specific parametric assumptions on the marginal distributions to proceed with estimation(e.g. Natarajan et al. [2009], Mishra et al. [2014] and Yan et al. [2022]).Our formulation provides the first procedure to obtain an MDM with best fit to choice data nonparametrically, while utilizing grouping information available (if any).
- We utilize the feasibility conditions developed to develop novel prediction intervals for choice probabilities for assortments unseen in past data.

Reference

V. K. Mishra, K. Natarajan, D. Padmanabhan, C.-P. Teo, and X. Li. On theoretical and empirical aspects of marginal distribution choice models. Management Science, 60(6):1511–1531, 2014.

K. Natarajan, M. Song, and C.-P. Teo. Persistency model and its applications in choice modeling. Management Science, 55(3):453–469, 2009.

Z. Yan, K. Natarajan, C.-P. Teo, and C. Cheng. A representative consumer model in data-driven multi-product pricing optimization. To appear in Management Science, 2022.