A Generalized Markov Chain Model to Capture Dynamic Preferences and Choice Overload

Kumar Goutam * Vineet Goyal † Agathe Soret ‡

Abstract

Assortment optimization is an important problem that arises in many practical applications such as retailing and online advertising where the goal is to find a subset of products from a universe of substitutable products that maximize a seller’s expected revenue. The demand and the revenue depend on the substitution behavior of the customers that is captured by a choice model. One of the key challenges is to find the right model for the customer substitution behavior. Many parametric random utility based models have been considered in the literature to capture substitution. However, in all these models, the probability of purchase increases as we add more options to the assortment. This is not true in general and in many settings, the probability of purchase may decrease if we add more products to the assortment, referred to as the choice overload. In this paper we attempt to address these serious limitations and propose a generalization of the Markov chain based choice model considered in Blanchet et al. [2]. In particular, we handle dynamic preferences and the choice overload phenomenon using a Markovian comparison model that is a generalization of the Markovian substitution framework of [2]. The Markovian comparison framework allows us to implicitly model the search cost in the choice process and thereby, modeling both dynamic preferences as well as the choice overload phenomenon. We consider the assortment optimization problem for the cases of our generalized Markov chain model when the underlying Markov chain is rank-1 (this is a generalization of the Multinomial Logit model) and when it has a low rank (this is a generalization of the mixture of MNLs model). We show that the assortment optimization problem under these models is NP-hard and also present a fully polynomial-time approximation scheme (FPTAS) in both these cases.

Key words: Assortment optimization, choice overload, choice model, Markov chain

*Industrial Engineering and Operations Research, Columbia University, New York, NY. kg2621@columbia.edu
†Industrial Engineering and Operations Research, Columbia University, New York, NY. vgoyal@ieor.columbia.edu
‡Industrial Engineering and Operations Research, Columbia University, New York, NY. acs2298@columbia.edu
1 Introduction

Assortment optimization problems arise in many practical applications such as retailing or online advertising. In such settings, a seller or decision maker has to select a subset of products from a universe of substitutable products to offer to customers in order to maximize the expected revenue. The demand of any item, and therefore the expected revenue, depend on the substitution behavior of the customers. Hence the importance of choosing the "right choice model" which specifies the probability that a random customer decides to buy a particular item offered in the set. The objectives are twofold: first determine or learn how customers choose and substitute among products, and second develop algorithms to find the optimal assortment.

More specifically, suppose we are given a universe of $n$ substitutable products, $N = \{1, ..., n\}$ with prices $p_1, ..., p_n$. Let $\pi(i, S)$ denote the probability that a random buyer selects product $i$ when the set of offered products is $S \subseteq N$. This probability depends on the substitution behavior of customers and is referred to as the choice probability. The assortment optimization problem, where goal is to find the subset $S$ of products that maximizes the expected revenue, can be formulated as

$$\max_{S \subseteq N} \sum_{i \in S} \pi(i, S)p_i$$

Many parametric choice models have been widely studied in the literature to capture customer substitution behavior and the probability $\pi(i, S)$. Most choice models are based on a random utility framework where the utility of product $j$ for any customer is $\tilde{u}_j = u_j + \xi_j$, where $u_j$ is the deterministic utility depending on product attributes and $\xi_j$ is the random component that also captures the idiosyncratic customer choice.

A choice model is specified by the choice of the deterministic utility $u_j$ and the distribution of $\xi_j$ and $\pi(j, S) = P(\tilde{u}_j \geq u_i, \forall i \in S \cup \{0\})$ where 0 is the no-purchase option.

One of the most popular choice model in practice is the Multinomial Logit (MNL) model. The MNL model was introduced independently by Luce [13] and Plackett [16], but came to be known as the MNL model after McFadden [14]. In this model, the random component of the utility is assumed to be i.i.d. according to the standard Gumbel distribution. For the MNL model, the probability that a customer purchases the product $i \in S$ when the subset $S$ is offered is given by

$$\pi(i, S) = \frac{e^{u_i}}{\sum_{j \in S} e^{u_j} + e^{u_0}} = \frac{v_i}{\sum_{j \in S} v_j + v_0},$$

where $v_j = e^{u_j}$. Talluri and van Ryzin [19] show that both estimation and assortment optimization under this model are tractable. More precisely, the optimal assortment is nested by price order, i.e. is composed of the top $k$ priced products for some $k \leq n$. Several algorithms including greedy, local search and linear programming based methods are known to solve the assortment optimization problem under the MNL model (see Talluri and van Ryzin [19], Gallego et al. [9], Davis et al. [4], and Jagabathula [11]).
However, this model suffers from certain simplifying assumptions, such as the Independence of Irrelevant Alternatives property (Ben-Akiva and Lerman [1]), which limit its applicability in many practical settings.

Therefore, more complex models have been developed to capture a richer class of substitution behaviors including the Nested Logit Model and the mixture of Multinomial Logit model. In the Nested Logit model (Williams [23]) when products are partitioned into nests. However, both estimation and assortment optimization become more challenging under this model. For instance, the assortment optimization problem under this model is NP-hard in general [5]. Davis et al. [5] show that under specific assumptions this problem is polynomially solvable.

In the mixture of Multinomial Logit models, we consider the population to be a mixture of several segments, each of which is given by a MNL. The probability of selecting product $i \in S$ when the set $S$ is offered is given by

$$\pi(i, S) = \frac{\sum_{k=1}^{p} \alpha_k}{\sum_{j \in S} v_{ik} + v_{0k}}$$

where $p$ is the number of segments, $\alpha_k$ for all $k \in [p]$ denotes the probability that a random customer belongs to segment $k$ ($\alpha_1 + \ldots + \alpha_p = 1$), and $v_{ik} \in \mathbb{R}^+$ for all $k \in [p]$ denote the MNL parameters for segment $k$. However, even for a mixture of MNL model with two segments, Rusmevichientong et al. [17] show that the assortment optimization problem is NP-hard. Moreover, Désir et al. [7] show that it is hard to approximate within a factor better than $\Omega(n^{1-\epsilon})$ in general. So the mixture of MNL model is quite intractable.

In addition, ranking-based choice models of demand have been studied in the literature (see Jagabathula and Rusmevichientong [11], Farias et al. [8], Désir et al. [6], [12], [20]). With the whole spectrum of choice models, finding the right model to capture the customer preferences in a particular application is a challenging problem. This is especially true since we only observe sales and not the complete preferences of the customers. Blanchet et al. [2] address the model selection problem and present a Markov chain based model where the substitutions are modeled via transitions in a Markov chain. They show that the Markov chain model captures the Multinomial Logit model exactly and provides a good approximation for any random utility based model under some fairly general assumptions. Furthermore, they show that the assortment optimization can be solved efficiently; thereby, providing good balance between predictive power and tractability. The Markov chain based model was first considered in Zhang and Cooper [24] and several variants of this model have since been considered (see for instance [3] and [15]).

However, all the random utility based models and distribution over ranking models suffer from two serious limitations in practice, namely, i) the models assume that customer preferences are static and exogenous to the set of products offered by the seller, and ii) the total probability of buying any product always increases (not necessarily strictly) if the seller adds more products to the assortment. In many settings, these properties are not satisfied. In particular, the customers may form their preferences based on the offered set of products and also, the purchase probability might decrease when the seller adds
more products to the assortment. This is referred to as the choice overload phenomenon and has been observed empirically in practice (see [10] and [18]). None of the random utility models capture this. Sahin and Wang [21] propose a model that incorporates an explicit search cost for users and can capture choice overload in some settings. Wang and Wang [22] also consider a choice model with endogenous network effects that capture dynamic preferences in some settings, mainly, where the utility of product for a customer depends on the number of customers interested in that product.

1.1 Our contributions

The main goal of this paper is to develop a model for substitutions that addresses the above limitations and lead to a more practical framework for choice modeling and assortment optimization. We propose a generalization of the Markov chain model introduced in [2] where we consider a Markovian comparison based choice process instead of one that is only based on Markovian substitutions.

Generalized Markov chain Model. We consider a generalization of the Markov chain model in [2], we consider a Markovian model to capture preferences. As in [2], we consider \((n + 1)\) states, one corresponding to each of the \(n\) products, and a state corresponding to the no-purchase option. Each customer arrives in a state corresponding to his or her most preferred item and transitions according to the transition probabilities, \(\rho_{ij}\) for all \(i = 1, \ldots, n\), \(j = 0, 1, \ldots, n\). If the customer reaches a state corresponding to a product offered in the assortment, she/he selects that product with some probability (possibly less than 1 unlike in [2]) that depends on the remaining products offered in the assortment.

More specifically, let \(\mu(i, S)\) denote the probability of selecting product \(i \in S\) given that the customer is in state \(i\) of the Markov chain. We model \(\mu(i, S)\) as a decreasing function of \(\sum_{j \in S} \rho_{ij}\) which can be interpreted as a measure of number of products in \(S\) similar to \(i\). Therefore, if a large number of similar products are offered in the assortment, the probability of selecting \(i\) given the customer is on state \(i\) is small. Intuitively, this can be interpreted as the higher search cost for finding the best products if many similar items are offered. In particular, we consider the following form of the stopping probability, \(\mu(i, S)\):

\[
\mu(i, S) = \exp \left( -\alpha \cdot \sum_{j \in S \cup \{0\}} \rho_{ij} \right),
\]

where \(\alpha \geq 0\). This probability is strictly less than one for \(\alpha > 0\) unlike the Markov chain model in [2]). We refer to this model as the Generalized Markov Chain Model. We would like to note that our model generalizes the model introduced in [2]. In particular, we recover the model in [2] by assuming \(\alpha = 0\).

Our model attempts to address the limitations of the random utility and rank-based choice models mentioned earlier in the following sense.

- **Dynamic Preferences.** Our model captures dynamic preferences of the customers. In other words, the substitution behavior of the customers in our model can be different for different assortments.
More specifically, the implied distribution over preferences depends on the assortment offered and there may not be any single distribution over preferences that is consistent with choices for all assortments. To the best of our knowledge, this is the first systematic approach to captures dynamic preferences.

- **Choice Overload Phenomenon.** An important consequence of our model is capturing the choice overload phenomenon. In particular, the probability of purchase does not necessarily increase if the seller includes more products in the assortment. More specifically, consider assortments $S, T \subseteq \{1, \ldots, n\}$ such that $S \subseteq T$. Then it is not necessarily true that $\pi(0, S) \geq \pi(0, T)$, where $\pi(0, S)$ denotes the probability of no purchase when the set $S$ of products is offered. We present several examples illustrating this.

**Generalized Multinomial Logit Model.** We consider the special case of the above generalized Markov chain model where the underlying Markov chain has rank one. Blanchet et al. [2] show that the Multinomial Logit model can be exactly captured by a Markov chain model where the transition probability matrix has rank one. In particular, suppose the MNL model is given by parameters $v_0, v_1, \ldots, v_n$. Without loss of generality, we can assume that $v_0 + \ldots + v_n = 1$. The equivalent Markov chain model has the following parameters:

- arrival probability to state $i$, $\lambda_i = v_i$, $\forall i = 0, \ldots, n$
- transition probabilities, $\rho_{ij} = v_j$, $\forall i = 1, \ldots, n$, $j = 0, \ldots, n$.

In our generalized model, we consider the stopping probability on state $i$ given that the customer is in state $i$, $\mu(i, S)$. We refer to this model as the **Generalized Multinomial Logit model.** We present several examples to illustrate that the above model captures dynamic preferences and the choice overload phenomenon.

We consider the assortment optimization problem under the above Generalized MNL model and show that it is NP-hard by a reduction from the partition problem. On the positive side, we present a fully polynomial time approximation scheme (FPTAS) for the assortment optimization problem.

Our algorithm is based on exploiting the structure of the choice probability function and consequently the revenue function. In particular, we show that the choice probability function for any assortment exhibits a nice rational functional form. While the revenue maximization problem is non-convex, we show that we can obtain a convex approximation of the objective function by guessing the values of a small number of linear functions for the optimal assortment. Our algorithm is a dynamic programming based algorithm that adapts ideas from the dynamic programming algorithm for the knapsack problem. Furthermore, our approach extends where there are a constant number of capacity or budget constraints on the assortments. These are natural constraints that arise commonly in practice.
Outline. The rest of the paper is organized as follows. In Section 2, we present the Generalized Markov Chain model and our notations, and we discuss some examples. In Section 3, we present the properties of the Generalized Multinomial Logit model, a special case of the previous model, including choice probability computations. In Section 4, we present the NP-hardness of the assortment optimization problem for the Generalized MNL model, and in Section 5, we present a fully polynomial-time approximation scheme (FPTAS) for the assortment optimization problem. In Section 6, we present some results on another special case of the Generalized Markov chain model, where the initial transition matrix is of low rank. These results will allow us to prove a FPTAS for this model as well.

2 Generalized Markov Chain Model and Notations

In this section we present the Generalized Markov Chain Model and the notations that we will use for the rest of the paper. We assume that we are given a universe of \( n \) substitutable products indexed from 1 to \( n \): \( N = \{1, \ldots, n\} \). Given this set, we first construct a directed graph \( G \) and give the customers' substitution behavior on this graph.

2.1 Customer substitution behavior in the Markov chain model

We model the customer substitution behavior using transitions in a Markov chain on \((n+1)\) states, where there is a state corresponding to each product and a state for the no-purchase alternative.

We first describe the Markov chain choice model introduced in Blanchet et al. [2]. In this Markov chain model, the customer substitutions are modeled using a Markov chain over \((n+1)\) states, \( N = \{0, 1, \ldots, n\} \): there is one state for each of the \( n \) substitutable products and a state 0 for the no-purchase alternative. Let \( S \subseteq N \) be a subset of offered products, let \( S_+ := S \cup \{0\} \). The model is specified by the parameters \( \lambda_i \), \( i \in [n] \) and \( \rho_{ij} \), for all \( i \in [n] \) and \( j \in \{0, \ldots, n\} \).

- \( \lambda_i \) denotes the arrival probability at state \( i \): a customer enters the graph \( G \) with an arrival probability \( \lambda_i \) at the vertex \( i \),
- \( \rho_{ij} \) denotes the transition probability from state \( i \) to state \( j \).

In this model, when \( S \) is offered, all the states in \( S \) are absorbing: if the random of any customer reaches state \( i \in S \), then he/she selects product \( i \) with probability one. However, the assumption that the customer will buy product \( i \in S \) with probability one if he reaches it, implies that the model suffers from a certain limitation. Indeed in this model, the customer preferences do not depend on the offered set. Therefore, if the seller offers more products, the customer will more likely reach a state in the offered set, thus decreasing the probability of reaching the no-purchase alternative. But as we pointed out above, this is not true in practice. In practice, when there are too many options, it is more difficult to make a
decision, and therefore it is more probable to choose the no-purchase alternative. This is why we consider a Markovian model that captures customers’ preferences.

2.2 Substitution behavior in our model

In our model, we use the Markovian framework as above to model substitution behavior. However, we introduce a stopping probability function $\mu(i, S)$.

**Stopping probability function** For any $i \in S$, $\mu(i, S)$ denotes the probability that a customer selects product $i$ given that she/he is currently in state $i$ of the Markov chain. In the model considered by Blanchet et al. [2], this probability is equal to 1. In this paper, we aim to capture the following fundamental component of customer choice, namely, that customer preferences and eventual selection depend on comparisons between the offered products. To capture this behavior, we consider the following formulation for $\mu(i, S)$

$$\mu(i, S) = \exp \left( -\alpha \sum_{j \in N_+} \rho_{ij} x_j \right),$$

where $x_j = 1$ if and only if $j \in S$. Note that if a large number of products similar to $i$ (i.e. with large $\rho_{ij}$) are offered in the assortment, then the stopping probability is small. This reflects the scenario that it is difficult to select a product if a large number of similar options are available. Similarly, if we include more products in the assortment, $\mu(i, S)$ decreases. This models the fact that customers need more comparison and time (transitions) to select the best product.

Here $\alpha$ is a scale parameter that amplifies the comparison effect. We suppose that $\alpha \geq 0$. A large value of $\alpha$ implies a very picky and risk-averse customer. $\alpha = 0$ reduces to Blanchet et al. [2].

Note that the choice of exponential function is arbitrary. We use this as it is a parsimonious choice to model transition probability. But any other function that is decreasing in $\sum_{j \in N_+} \rho_{ij} x_j$ with range in $[0, 1]$ could have worked as well.

To model the fact that a customer has finally decided to purchase the product $i$, we add a vertex $i'$, which is an absorbing state. A directed edge joins the vertex $i$ to the vertex $i'$ with weight $\mu(i, S)$, which represents the probability of buying the product $i$ when the customer is at the vertex $i$. This probability is equal to 0 if and only if the product is not offered, i.e. $i \notin S$. We note $N'_+$ the set of absorbing states $\{i'\}_{i \in [n]} \cup \{0\}$. At a certain time $t$, for $t$ large enough, the customer is either in a certain state $i'$, with $i \in S$, or in the no-purchase state 0.

**Modified transition probabilities** Since the sum of the probabilities of getting out of $i$ has to be equal to 1, we change $\rho_{ij}$ to $\tilde{\rho}_{ij}$ defined as follows:

$$\tilde{\rho}_{ij} = (1 - \mu(i, S)) \rho_{ij}.$$  

**Customer substitution behavior on the graph** Let us summarize how a customer behaves on the graph
given a certain set of products $S \subseteq N$ to sell under the Generalized Markov Chain model:

- The customer arrives with probability $\lambda_i$ at the vertex $i$.
- If the product $i$ is in $S$ then the customer, currently at the vertex $i$,
  - either selects it with probability $\mu(i, S)$, arrives at the vertex $i'$ and then stops,
  - or goes to another vertex $j$ with probability $\bar{\rho}_{ij}$.
- If $i \notin S$, then the customer cannot purchase the product, so with probability $\bar{\rho}_{ij}$ he goes to another vertex $j$.
- If $i = 0$ then the customer has decided not to purchase any product, and he stops.

We then proceed recursively.

**Example** We consider the following 4-vertex graph (see Figure 1), where we have chosen to offer the subset $S = \{3, 4\}$.

![Figure 1: Example of a 4-vertex graph with $S = \{3, 4\}$](image)

### 2.3 Assortment Optimization Problem

Let $\pi(i, S)$ be the probability of buying the product $i$ when the subset $S$ is offered, and $p_i$ be the price of the product $i$. We finally assume that the seller has a fixed capacity of goods, therefore the cost price does not depend on our decision. The assortment optimization is the following

$$\max_{S \subseteq N} \sum_{i \in S} \pi(i, S) \times p_i.$$ 

Now the objective is to compute $\pi(i, S)$ under our model.
3 Computation of Choice Probabilities

Given the parameters \( \lambda_i, \bar{\rho}_{ij} \) and \( \mu(i, S) \) for all \( i \in N \) and \( j \in N_+ \), let us describe how we compute the choice probabilities for any \( S \subseteq N \). Our assumption is that a customer arrives at the state \( i \) with probability \( \lambda_i \), and continues to transition according to probabilities \( \bar{\rho}_{ij} \) until he decides to buy a product \( i \) with probability \( \mu(i, S) \) when he is at the vertex \( i \), or decides not to buy any product and ends at the no-purchase vertex 0. We therefore assume that any customer buys at most one product.

3.1 Choice probabilities

Let \( \rho(N, N) \) be the transition probability matrix from states \( N \) to \( N \). We recall that since the probability of exiting a vertex \( i \) is 1, and since there is a probability \( \mu(i, S) \) of buying the product represented by the vertex, the transition probabilities are given by: \( \bar{\rho}_{ij} = (1 - \mu(i, S)) \rho_{ij} \). Consequently, \( \rho(N, N) = Diag((1 - \mu(i, S))) \times \rho \),

where \( \rho \) is the initial transition probability matrix, \( \rho = (\rho_{ij})_{i,j \in [n]} \), and \( Diag((1 - \mu(i, S))) \) is the diagonal matrix with \((1 - \mu(i, S))\) on its diagonal.

After a certain time, every customer will be in an absorbing state. In order to compute \( \pi(i, S) \), we have to know the probability that a customer arrives at the vertex \( i' \). For \( i \in [1, n] \) we have:

\[
\pi(i, S) = \lim_{q \to \infty} \lambda^T(\mathcal{P}(S))^q e^{2n+1}_i,
\]

where \( \mathcal{P}(S) \) is the transition probability matrix in the graph when the subset \( S \) is offered and is of the form:

\[
\mathcal{P}(S) = \begin{bmatrix}
\rho(N'_+, N'_+) & \rho(N'_+, N) \\
\rho(N, N'_+) & \rho(N, N)
\end{bmatrix} = \begin{bmatrix}
I_{n+1} & 0 \\
\Pi(S) & Diag((1 - \mu(i, S)))\rho
\end{bmatrix},
\]

and \( \Pi(S) \) is the following matrix:

\[
\Pi(S) = \rho(N, N'_+) = \begin{bmatrix}
\mu(1, S) & 0 & \ldots & 0 & \bar{\rho}_{10} \\
0 & \mu(2, S) & \ldots & 0 & \bar{\rho}_{20} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \ldots & 0 & \mu(n, S) & \bar{\rho}_{n0}
\end{bmatrix} = \begin{bmatrix}
Diag(\mu(i, S)) & \bar{\rho}_0
\end{bmatrix},
\]

and \( e^{2n+1}_i \in \{0, 1\}^{2n+1} \) has 0 on each component except on his \( i^{th} \) component. Since we take \( i \in [1, n] \), \( e^{2n+1}_i \) will always have his \( n + 1 \) last components equal to 0. \( \rho(N'_+, N'_+) = I_{n+1} \) because all the states in \( N'_+ \) are absorbing, which also implies that \( \rho(N'_+, N) = 0 \).
For \( q \in \mathbb{N} \), we have:

\[
P(S)^q = \begin{bmatrix} I_{n+1} & 0 \\ \sum_{k=0}^{q} (\text{Diag}((1 - \mu(i, S)))\rho)^k \Pi(S) & (\text{Diag}((1 - \mu(i, S)))\rho)^q \end{bmatrix}.
\]

Therefore, if we assume that the spectral radius of \( \rho(N, N) = \text{Diag}((1 - \mu(i, S)))\rho \) is strictly less than 1:

\[
\lim_{q \to \infty} P(S)^q = \begin{bmatrix} I_{n+1} & 0 \\ (I_n - \text{Diag}((1 - \mu(i, S)))\rho)^{-1} \Pi(S) & 0 \end{bmatrix}.
\]

And

\[
\pi(i, S) = \lambda^T (I_n - \text{Diag}((1 - \mu(i, S)))\rho)^{-1} \Pi(S)e_i,
\]

where \( e_i \in \{0, 1\}^{n+1} \). If we want the probability of no purchase, we will compute \( \lambda^T (I_n - \text{Diag}((1 - \mu(i, S)))\rho)^{-1} \Pi(S)e_0 \), where \( e_0 = (0, ..., 0, 1) \in \{0, 1\}^{n+1} \).

### 3.2 Examples

We provide now some examples to show that our model is better at capturing the choice overload phenomenon than any other random utility based models.

**Example 1 (Homogeneous graph).** We first consider the case of the complete graph with \( n \) vertices with homogeneous transition probabilities, \( \rho_{ij} = \frac{1}{n+1} \) for all \( i \in N \) and \( j \in N_+ \), and homogeneous probabilities of arrival, \( \lambda_i = \frac{1}{n+1} \) for all \( i \in N_+ \). We suppose that all the products have the same price \( p \). The symmetry of this example implies that we only need to find the number \( k \) of vertices to offer that would maximize the revenue, and then take randomly a subset of \( k \) elements. For any random utility based choice models presented in the first section, the optimal set to maximize our revenue will be the entire set of products. Indeed, in all these models, the product with the highest price is always in the optimal assortment. Since all the products have the same price, they will all be in the set. Furthermore, the probability of no purchase always decreases when we add more products into the offered set. However this is not true in our model, and the probability of no purchase depends on the parameter \( \alpha \) chosen. We have chosen a universe of \( n = 15 \) substitutable products, and given the assumptions above, we compute the probability of purchase under our model when \( k \) products are in the offered set, for all \( k \in [n] \) and we obtain the following graph:

We see that, depending on the value of \( \alpha \), the probability of no purchase may increase when we add more items into the offered assortment. Especially, a high \( \alpha \) implies a sooner (in terms of the number of products) increase in the probability of no-purchase.

We present another example, the star graph, which shows that our model will favour clusters’ centers as items for the assortment set, unlike the Markov chain model.
Example 2 (Star Graph). We consider the following star graph with $n$ vertices. We suppose that the vertex 1 is linked to all the other vertices in the graph, but that the other vertices are only linked to 1 and to the no purchase vertex 0. We suppose that the transition probabilities are homogeneous. Therefore, considering the structure of the graph, they are given by:

$$
\rho_{1i} = \frac{1}{n} \quad \forall i \in N \setminus \{1\}
$$

$$
\rho_{i1} = \rho_{i0} = \frac{1}{2} \quad \forall i \in N \setminus \{1\}.
$$

We suppose that the arrival probabilities are all equal $\lambda_i = \frac{1}{n}$ for all $i \in N$. Finally we suppose that all the products, except 1, have a price $P$, but that the product 1 has a smaller price $p < P$.

For any random utility based choice models considered in the literature, since the product with the highest price is always in the optimal offered set, then $\{2, ..., n\} \subseteq S^*$ where $S^*$ is the optimal set. And this is true, for any $n$, even for very large $n$.

However, our model considers that selling only the product 1 will give a higher revenue, when $\alpha$ is large enough. For $\alpha \geq 9$, the optimal set will always be $\{1\}$. And this result is closer to the reality. Indeed, it seems more logical in practice for the seller to only offer the product which is similar to many other products, even if this product is slightly less expensive than the others.

In order to get a generalization of the MNL model, we will now suppose that the initial transition probability matrix, $\rho = (\rho_{ij})_{i,j \in [n]}$ is of rank one, and show that with such an assumption, the optimization problem in NP-hard.
4 Generalized Multinomial Logit model

In this section we suppose that the transition probability matrix is of rank one. Given this assumption, we refer to our model as the Generalized Multinomial Logit model, which is a special case of the Generalized Markov chain model. We remind that \( N \) represents all the vertices in the graph, \( N_+ \) is the union of all the vertices and the vertex 0 which represents the case where the customer does not buy a product, and \( \rho(N, N_+) \) is the transition probability matrix from \( N \) to \( N_+ \). In this model, we suppose that there exists \( v = (v_i)_{i\in [n+1]} \in [0,1]^{n+1} \) such that \( \sum_{i=0}^{n} v_i = 1 \) and:

\[
\rho(N, N_+) = \begin{bmatrix}
1 - \mu(1, S) & 0 & \cdots & 0 \\
0 & 1 - \mu(n, S)
\end{bmatrix} \times \begin{bmatrix}
v_0 & v_1 & \cdots & v_n
\end{bmatrix}.
\]

Since there is a probability \( \mu(i, S) \) that the customer goes from vertex \( i \) to vertex \( i' \), the probability that the customer goes from vertex \( i \) to vertex \( j \), with \( j \in N_+ \), has to be \( (1 - \mu(i, S))v_j \), therefore the probability of exiting from vertex \( i \) is 1. We also suppose in this model that \( \lambda = v \), therefore the probability of arriving at a vertex \( i \) is proportional to \( v_i \). Finally we suppose that the probability of buying the product \( i \) while being at vertex \( i \) is given by:

\[
\mu(i, S) := e^{-\alpha \sum_{j \in S^+} v_j}.
\]

With the notations given in Section 2:

\[
\mathcal{P}(S) = \begin{bmatrix}
\rho(N_+^I, N_+^I) & \rho(N_+^I, N) \\
\rho(N, N_+^I) & \rho(N, N)
\end{bmatrix} = \begin{bmatrix}
I_{n+1} & 0 \\
\Pi(S) & D(S)\rho_v
\end{bmatrix},
\]

where

\[
\Pi(S) = e^{-\alpha \sum_{j \in S^+} v_j} \times \begin{bmatrix}
\mathbb{1}_{1 \in S} & 0 & (e^{\alpha \sum_{j \in S^+} v_j} - \mathbb{1}_{1 \in S})v_0 \\
\vdots & \ddots & \vdots \\
0 & \mathbb{1}_{n \in S} & (e^{\alpha \sum_{j \in S^+} v_j} - \mathbb{1}_{n \in S})v_0
\end{bmatrix},
\]

and

\[
D(S)\rho_v = \begin{bmatrix}
1 - \mu(1, S) & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & 1 - \mu(n, S)
\end{bmatrix} \times \begin{bmatrix}
v_1 & \cdots & v_n
\end{bmatrix} = (1 - e^{-\alpha \sum_{j \in S^+} v_j}) \times \begin{bmatrix}
v_1 & \cdots & v_n
\end{bmatrix}.
\]

We summarize here the assumptions of the Generalized Multinomial Logit model.

**Generalized Multinomial Logit model** In this model we make the following assumptions:
• the initial transition probability matrix within \( N_+ \) is of rank one, i.e. there exists \( v = (v_i)_{i \in N_+} = \left( \begin{array}{c} v_0 \\ v_* \end{array} \right) \in [0,1]^{n+1} \) such that

\[
\rho(N,N) = \text{Diag}((1 - \mu(i,S))1v_*^T \quad \text{and} \quad \sum_{j \in N} v_j + v_0 = 1 ,
\]

• given a subset \( S \subseteq N \), for all \( i \in N \) we have \( \mu(i,S) = e^{-\alpha \sum_{j \in S_+} v_j} \)

• and for all \( j \in N_+ \), \( \lambda_j = v_j \).

As we showed in Section 2, our assortment optimization problem under our model is given by:

\[
\max_{S \subseteq N} \lambda^T(I_n - D(S)\rho_v)^{-1}\Pi(S)p,
\]

and we will now give an exact formulation of the probability of buying a product and see why it is a generalization of the MNL model.

4.1 Choice probability

We can compute the probability of buying a product \( i \) in our model as follows.

**Lemma 4.1.** The probability of buying a product \( i \) given a chosen subset \( S \subseteq N \) under the Generalized Multinomial Logit model is:

\[
\pi(i,S) = \frac{v_i}{\sum_{k \in S} v_k + v_0 e^{\alpha \sum_{j \in S_+} v_j}} 1_{i \in S}.
\]

**Proof.** Let \( S \subseteq N \), and \( (I_n - D(S)\rho_v)^{-1} = (x_{ij})_{i,j \in N} \), our assortment optimization problem becomes:

\[
\max_{S \subseteq N} \lambda^T(I_n - D(S)\rho_v)^{-1}\Pi(S)p = \sum_{i \in S} \left( \sum_{k \in N} \lambda_k x_{ki} \right) e^{-\alpha \sum_{j \in S_+} v_j} p_i.
\]

Let \( i \in S \) we want to compute

\[
\pi(i,S) = \left( \sum_{k \in N} \lambda_k x_{ki} \right) e^{-\alpha \sum_{j \in S_+} v_j}.
\]

We can show that:

\[
\forall k,j \in N \quad \begin{cases} 
\frac{x_{kj}}{x_{j}(1 - \mu(k,S))} = \frac{1}{\sum_{s \in S} v_s \mu(s,S) + v_0} & \text{if } k \neq j, \\
\frac{x_{jk} - 1}{x_k(1 - \mu(k,S))} = \frac{1}{\sum_{s \in S} v_s \mu(s,S) + v_0} & \text{otherwise}.
\end{cases}
\]
Let $\pi_S = e^{-\alpha \sum_{j \in S^+} v_j}$ and $x = \frac{1}{\sum_{k \in N} v_k \mu(k, S) + v_0} = \frac{1}{\pi_S \sum_{k \in S} v_k + v_0}$, then:

$$
\sum_{k \in N} \lambda_k x_{ki} = \lambda_i x_{ii} + \sum_{k \neq i} \lambda_k x_{ki} (1 - \mu(k, S))
$$

$$
= \lambda_i x_{ii} + x v_i \sum_{k \notin S} \lambda_k + x v_i (1 - \pi_S) \sum_{k \in S \setminus \{i\}} \lambda_k
$$

$$
= \lambda_i x_{ii} + x v_i (1 - \lambda_i - \lambda_0) - x v_i \pi_S \sum_{k \in S \setminus \{i\}} \lambda_k
$$

$$
= \lambda_i (1 + (1 - \pi_S) x v_i) + x v_i (1 - \lambda_i - \lambda_0) - x v_i \pi_S \sum_{k \in S \setminus \{i\}} \lambda_k
$$

$$
= \lambda_i + x v_i \left(1 - \lambda_0 - \pi_S \sum_{k \in S} \lambda_k\right)
$$

Since we supposed that $\lambda_j = v_j$ for all $j \in N$, then the probability of buying the product $i$ becomes:

$$
\pi(i, S) = \left(\sum_{k \in N} \lambda_k x_{ki}\right) e^{-\alpha \sum_{j \in S^+} v_j}
$$

$$
= \pi_S v_i \left(1 + \frac{1 - v_0 - \pi_S \sum_{k \in S} v_k}{\pi_S \sum_{k \in S} v_k + v_0}\right)
$$

$$
= \pi_S v_i \left(\frac{1}{\pi_S \sum_{k \in S} v_k + v_0}\right)
$$

$$
\pi(i, S) = \frac{v_i}{\sum_{k \in S} v_k + v_0 e^{\alpha \sum_{j \in S^+} v_j}}
$$

And if $i \notin S$, we have of course $\pi(i, S) = 0$ which finishes the proof.

**Our model is a generalization of the MNL model.** We recall that under the MNL model, the probability of buying the product $i \in S$ when the set $S$ is offered is

$$
\pi_{MNL}(i, S) = \frac{v_i}{\sum_{j \in S} v_j + v_0}.
$$

Therefore, our model can be considered as a generalization of the MNL model where the no-purchase probability is not a constant but depends on the assortment as $v_0 e^{\alpha \sum_{j \in S^+} v_j}$, which increases the utility of the no-purchase alternative in the MNL model.

Suppose that, instead of choosing, $\mu(i, S) = e^{-\alpha \sum_{j \in S^+} v_j}$, we had chosen $\mu(i, S) = \frac{1}{\sum_{j \in S^+} v_j}$, which would also convey the idea that $\mu(i, S)$ is a decreasing function of $\sum_{j \in S^+} v_j$. Then, the probability of purchasing the product $i$ given a set $S$ of offered products would have been:

$$
\bar{\pi}(i, S) = \frac{v_i}{\sum_{j \in S} v_j + v_0 \times \left(\sum_{j \in S^+} v_j\right)} = \frac{v_i}{(1 + v_0) \sum_{j \in S} v_j + v_0^2},
$$

which is in a ratio scale, exactly the choice probability of the MNL model. Therefore, such a function
would have given a nesting by fare order choice model, as the MNL, and therefore an optimization prob-
lem solvable in polynomial time. However, this model would not have given sufficient weights to the $v_j$’s
for $j \in S$ on the no-purchase alternative, which is what we want to capture the choice overload phe-
nomenon. This is why we want to emphasize the importance of the choice of the function $\mu(i, S)$ in our
model, and why the choice of $e^{-\alpha \sum_{j \in S} v_j}$ meets the requirements of our model.

4.2 Example

Let us revisit the example of a homogeneous Markov chain from Section 3 in light of the exact choice
probability computations.

Example (Homogeneous graph). We recall the assumptions in this model. consider the case of the
complete graph with $n$ vertices with homogeneous transition probabilities, $\rho_{ij} = \frac{1}{n+1}$ for all $i \in N$ and
$j \in N_+$, and homogeneous probabilities of arrival, $\lambda_i = \frac{1}{n+1}$ for all $i \in N_+$. We suppose that all the
products have the same price $p$.

For any random utility based choice models presented in the first section, the optimal set to maximize
our revenue will be the entire set of products, as we explained before. This example is a particular case
where the initial transition probability is of rank one. Therefore under the Generalized Multinomial Logit
model, the assortment optimization problem for the homogeneous graph is:

$$\max_{S \subseteq N} \frac{|S|}{n+1} p - \frac{|S|}{n+1} e^{\alpha \sum_{i \in S} v_i} = \max_{k \in [n]} kp - e^{\alpha \sum_{i \in S} v_i}.$$

A simple computation shows that the optimal number of products in the offered set is $k^* = \frac{n+1}{\alpha}$. There-
fore, if $\alpha < \frac{n+1}{n}$ then the optimal assortment set will be the entire as in the previous models. However,
if we take $\alpha$ large enough, then $\frac{n+1}{\alpha} \leq n - 1$ and there will be less products in the optimal offered set.

This also highlights the function or the meaning of $\alpha$: $\alpha$ amplifies the comparison effect. A large value
of $\alpha$ implies risk-averse customer, and therefore a strategy where the seller offers less products.

4.3 NP-hardness of the Assortment Optimization problem

Under the Generalized Multinomial Logit model, the optimization problem is

$$\max_{S \subseteq N} R(S) := \max_{S \subseteq N} \frac{\sum_{i \in S} v_i p_i}{\sum_{i \in S} v_i + \alpha \sum_{i \in S} v_i}.$$

In this section we prove that this problem is NP-hard, for $\alpha \leq 1$ and $\alpha > 2$. To show this, we first prove
the following structural property of an optimal solution, when $\alpha \leq 1$.

Lemma 4.2. In the Generalized Multinomial Logit model with $\alpha \leq 1$, the product with the highest price
we want to prove that $g$ is a strictly decreasing function on $(0, 1)$. Since $g(0) = 0$, we will make a reduction from the partition problem. Consider the following instance of the partition problem: we are given $n$ integers $c_1, ..., c_n$ and the goal is to decide whether there is a subset $S \subseteq \{1, ..., n\}$ such that $\sum_{i \in S} c_i = \sum_{i \in \{1, ..., n\} \setminus S} c_i$. Let $T = \frac{1}{2} \sum_{i=1}^n c_i$, then $\sum_{i \in S} c_i = \sum_{i \in \{1, ..., n\} \setminus S} c_i$ if and only if $\sum_{i \in S} c_i = T$. We can suppose

\[
R(S \cup \{1\}) \geq R(S).
\]

**Proof.** If $N \setminus \{1\} = \emptyset$ then the result is trivial since $R(\emptyset) = 0$. Now, suppose that there is at least one other product than 1 in $N$. Let $S \subseteq N \setminus \{1\}$. We will use the following notations:

\[
V(S) = \sum_{j \in S} v_j \quad \text{and} \quad V P(S) = \sum_{j \in S} v_j p_j.
\]

Therefore we have

\[
R(S \cup \{1\}) - R(S) \geq 0 \iff \frac{V P(S) + v_1 p_1}{V(S) + v_1 + v_0 e^{\alpha V(S)} e^{\alpha v_1}} - \frac{V P(S)}{V(S) + v_0 e^{\alpha V(S)}} \geq 0 \iff v_1 (p_1 V(S) - V P(S)) + v_0 e^{\alpha V(S)} (v_1 p_1 - (e^{\alpha v_1} - 1) V P(S)) \geq 0
\]

Since 1 has the highest price, $p_1 V(S) \geq V P(S)$ and

\[
(e^{\alpha v_1} - 1) V P(S) \leq (e^{\alpha v_1} - 1) V(S) p_1 = (e^{\alpha v_1} - 1)(1 - v_1 - \beta(S))p_1,
\]

where $\beta(S) = \sum_{k \in N \setminus \{1\} \cup S} v_k$. Moreover, since

\[
(e^{\alpha v_1} - 1)(1 - v_1 - \beta(S)) - v_1 = e^{\alpha v_1} (1 - v_1 - \beta(S)) - 1 + \beta(S),
\]

we want to prove that $g : x \mapsto e^{\alpha x} (1 - x - \beta(S)) - 1$ is a negative function on $(0, 1)$. Indeed, it is a strictly decreasing function on $(0, 1)$:

\[
g'(x) = e^{\alpha x} (\alpha - \alpha x - \alpha \beta(S) - 1) < 0 \iff 1 - \beta(S) - x < \frac{1}{\alpha}.
\]

And $1 - \beta(S) - x = V(S) \leq 1 - v_0 < \frac{1}{\alpha}$ since we assumed $\alpha \leq 1$. Moreover, $g(0) = 0$, therefore $g$ is negative on $(0, 1)$ and we have $R(S \cup \{1\}) - R(S) \geq 0$.

We can finally prove the main theorem of this section.

**Theorem 4.3.** The optimization problem under the Generalized Multinomial Logit model is NP-hard for $\alpha \leq 1$.

**Proof.** To prove this result we will make a reduction from the partition problem. Consider the following instance of the partition problem: we are given $n$ integers $c_1, ..., c_n$ and the goal is to decide whether there is a subset $S \subseteq \{1, ..., n\}$ such that $\sum_{i \in S} c_i = \sum_{i \in \{1, ..., n\} \setminus S} c_i$.

Let $T = \frac{1}{2} \sum_{i=1}^n c_i$, then $\sum_{i \in S} c_i = \sum_{i \in \{1, ..., n\} \setminus S} c_i$ if and only if $\sum_{i \in S} c_i = T$. We can suppose
without loss of generality that $c_i > 0$ for all $i \in [n]$. We construct an instance of our problem as follows:

$$
v_i = \begin{cases} \frac{c_i}{2T + 1} & \text{if } i \geq 1, \\ 1 - \sum_{i=1}^n v_i = \frac{1}{2T + 1} & \text{if } i = 0. \end{cases}
$$

and let $c_0 := v_0 e^{\alpha v_0} > 0$, we define the prices as follows

$$
p_i = \begin{cases} \frac{1}{(2T + 1)c_0} + \frac{e^{\alpha T} - 1}{T} + \frac{1}{c_i} & \text{if } i = 1, \\ \frac{1}{(2T + 1)c_0} + \frac{e^{\alpha T} - 1}{T} & \text{otherwise.} \end{cases}
$$

Finally we set the target revenue as $K = \frac{T}{(2T + 1)c_0} + \frac{e^{\alpha T}}{T + (2T + 1)c_0 e^{\alpha T}}$.

First, we can note that 1 is necessarily in the optimal set. Indeed 1 has the highest price in $N$, so the previous lemma implies that 1 is necessarily in the optimal set (we can note that the choice of 1 is random and we could have chosen any $i$ in $N$).

In this case, our problem becomes

$$
\max_{S \subseteq \{1\ldots,n\}} R(S) := \max_{S \subseteq \{1\ldots,n\}} R(S \cup \{1\})
= \max_{S \subseteq \{2\ldots,n\}} \frac{\sum_{i \in S} v_i p_i + v_1 p_1}{\sum_{i \in S \cup \{1\}} v_i + c_0 e^{\alpha} \sum_{i \in S \cup \{1\}} v_i} 
= \max_{S \subseteq \{1\ldots,n\}} \frac{1}{(2T + 1)c_0} + \frac{e^{\alpha T} - 1}{T} \sum_{i \in S} c_i + 1
= \max_{S \subseteq \{1\ldots,n\}} F \left( \sum_{i \in S} c_i \right)
$$

where

$$
F : \left[0, 2T \right] \to R_+ \quad \xrightarrow{x} \quad \frac{h(x)}{x + (2T + 1)c_0 e^{\alpha T}},
$$

and

$$
h : x \mapsto \left( \frac{1}{(2T + 1)c_0} + \frac{e^{\alpha T} - 1}{T} \right) x + 1.
$$
Proof. \( \alpha > 2 \)

Theorem 4.4. The optimization problem under the Generalized Multinomial Logit model is NP-hard for instance of the partition problem: we are given \( \alpha > 2 \) \( n \) integers \( c_1, \ldots, c_n \) and the goal is to decide whether there is a subset \( S \subseteq \{1, \ldots, n\} \) such that \( \sum_{i \in S} c_i = T \). Therefore, since \( h(0) = g(0) = 1 \) and \( h(2T) < g(2T) \). Indeed,

\[
h(2T) - g(2T) = 2(e^{\frac{2T}{T+1}} - 1) + 1 - e^{\frac{2T}{T+1}} = -(e^{\frac{2T}{T+1}} - 1)^2 < 0.
\]

Therefore, since \( h \) is a line with a positive slope, there exists a unique \( x^* \in (0, 2T) \) such that for all \( 0 < x < x^* \), \( h(x) - g(x) > 0 \), \( h(x^*) - g(x^*) = 0 \) and for all \( 2T \geq x > x^* \), \( h(x) - g(x) < 0 \). But \( h(T) = g(T) \). So \( x^* = T \). So this proves that \( F \) is strictly increasing on \( [0, T) \) then strictly decreasing on \( (T, 2T] \). So \( F \) has a unique maximum at \( T \) on \( (0, 2T) \). Hence,

\[
\max_{S \subseteq \{1, \ldots, n\}} R(S) = \max_{S \subseteq \{1, \ldots, n\}} F \left( \sum_{i \in S} c_i \right) \leq F(T) = \frac{T}{c_0} + e^{\frac{\alpha T}{T+1}} = K.
\]

So there exists an assortment \( S \subseteq \{1, \ldots, n\} \) whose expected revenue is at least \( K \) if and only if the chain of inequalities hold as equalities. For this to happen we need to have \( \sum_{i \in S'} c_i = T \) for some assortment \( S' \subseteq \{1, \ldots, n\} \). Therefore there exists an assortment \( S \subseteq \{1, \ldots, n\} \) whose expected revenue is at least \( K \) if and only if there exists an assortment \( S' \subseteq \{1, \ldots, n\} \) that satisfies \( \sum_{i \in S'} c_i = T \). \( \square \)

Although when \( \alpha > 2 \), the product with the highest price is not necessarily in the optimal assortment set, we can also prove that the assortment optimization problem is NP-hard when \( \alpha > 2 \).

Theorem 4.4. The optimization problem under the Generalized Multinomial Logit model is NP-hard for
\( \alpha > 2 \).

Proof. To prove this result we will make a reduction from the partition problem. Consider the following instance of the partition problem: we are given \( n \) integers \( c_1, \ldots, c_n \) and the goal is to decide whether there is a subset \( S \subseteq \{1, \ldots, n\} \) such that \( \sum_{i \in S} c_i = \sum_{i \in \{1, \ldots, n\} \setminus S} c_i \).

Let \( T = \frac{1}{2} \sum_{i=1}^{n} c_i \), then \( \sum_{i \in S} c_i = \sum_{i \in \{1, \ldots, n\} \setminus S} c_i \) if and only if \( \sum_{i \in S} c_i = T \). We can suppose

\[
F'(x) \geq 0 \Leftrightarrow \frac{h'(x)(x + (2T + 1)c_0 e^{\frac{\alpha T}{T+1}}) - h(x)(1 + c_0 e^{\frac{\alpha T}{T+1}})}{(x + (2T + 1)c_0 e^{\frac{\alpha T}{T+1}})^2} \geq 0
\]

\[
\Leftrightarrow h'(x) - \frac{1 + c_0 e^{\frac{\alpha T}{T+1}}}{x + (2T + 1)c_0 e^{\frac{\alpha T}{T+1}}} h(x) \geq 0
\]

\[
\Leftrightarrow \frac{h'(x)}{h(x)} \geq \frac{1 + c_0 e^{\frac{\alpha T}{T+1}}}{x + (2T + 1)c_0 e^{\frac{\alpha T}{T+1}}} > 0 \text{ since } h > 0 \text{ on } [0, 1]
\]

\[
\Rightarrow \ln(h(x)) - \ln(h(0)) \geq \ln(x + (2T + 1)c_0 e^{\frac{\alpha T}{T+1}}) - \ln((2T + 1)c_0)
\]

\[
\Rightarrow h(x) \geq \frac{1}{(2T + 1)c_0} (x + (2T + 1)c_0 e^{\frac{\alpha T}{T+1}}) \text{ since } h(0) = 1
\]

\[
\Rightarrow h(x) - g(x) \geq 0
\]
without loss of generality that \( c_i > 0 \) for all \( i \in [n] \). We construct an instance of our problem as follows:

\[
v_i = \begin{cases} \frac{c_i}{T + \alpha} & \text{if } i \geq 1, \\ 1 - \sum_{i=1}^{n} v_i = 1 - \frac{2}{\alpha} & \text{if } i = 0, \end{cases}
\]

and let \( c_0 := v_0 e^{\alpha v_0} > 0 \). We note that \( v_0 > 0 \) because we have supposed \( \alpha > 2 \). We define the prices as follows

\[
\forall i \in [n] \quad p_i = 1.
\]

Finally we set the target revenue as \( K = \frac{1}{1 + \alpha c_0 e^{\alpha}} \). In this case, our problem becomes

\[
\max_{S \subseteq \{1, \ldots, n\}} R(S) := \max_{S \subseteq \{1, \ldots, n\}} \frac{\sum_{i \in S} v_i p_i}{\sum_{i \in S} v_i + c_0 e^{\alpha} \sum_{i \in S} v_i}
\]

\[
= \max_{S \subseteq \{1, \ldots, n\}} \frac{\frac{1}{T + \alpha} \left( \sum_{i \in S} c_i \right) + c_0 e^{\alpha} \sum_{i \in S} v_i}{\left( \sum_{i \in S} v_i \right) + c_0 e^{\alpha} \sum_{i \in S} v_i}
\]

\[
= \max_{S \subseteq \{1, \ldots, n\}} F \left( \sum_{i \in S} c_i \right),
\]

where

\[
F : [0, 2T] \to \mathbb{R}_+ \\
\quad x \mapsto \frac{x}{x + T \alpha c_0 e^{\alpha}}.
\]

\( F \) is increasing at \( x \) if and only if

\[
F'(x) \geq 0 \iff \frac{x + T \alpha c_0 e^{\alpha} - x(1 + \alpha c_0 e^{\alpha})}{(x + T \alpha c_0 e^{\alpha})^2} \geq 0
\]

\[
\iff x \leq T.
\]

Therefore \( F \) is strictly increasing on \([0, T]\) then strictly decreasing on \((T, 2T]\). So \( F \) has a unique maximum at \( T \) on \((0, 2T)\). Hence,

\[
\max_{S \subseteq \{1, \ldots, n\}} R(S) = \max_{S \subseteq \{1, \ldots, n\}} F \left( \sum_{i \in S} c_i \right) \leq F(T) = \frac{1}{1 + \alpha c_0 e^{\alpha}} = K.
\]

So there exists an assortment \( S \subseteq \{1, \ldots, n\} \) whose expected revenue is at least \( K \) if and only if the chain of inequalities hold as equalities. For this to happen we need to have \( \sum_{i \in S'} c_i = T \) for some assortment \( S' \subseteq \{1, \ldots, n\} \). Therefore there exists an assortment \( S \subseteq \{1, \ldots, n\} \) whose expected revenue is at least \( K \) if and only if there exists an assortment \( S' \subseteq \{1, \ldots, n\} \) that satisfies \( \sum_{i \in S'} c_i = T \). \( \square \)

Before we present an FPTAS for the assortment optimization problem in accordance with the hardness result we just proved, we discuss how to do a parameter estimation for this generalized MNL model from given data.
4.4 Parameter Estimation for Generalized MNL Model

The choice probabilities for the Generalized MNL model are given by:

\[ \pi(i, S) = \frac{v_i}{v_0 e^{\alpha \sum_{j \in S} v_j} + \sum_{k \in S} v_k}, \forall i \in S \]

\[ \pi(0, S) = \frac{v_0}{v_0 e^{\alpha \sum_{j \in S} v_j} + \sum_{k \in S} v_k}, \]

where \( v_j = e^{\beta^T x_j} \).

Given a choice dataset: \( D = \{j_t, S_t\}_{t=1}^T \), we have to estimate the parameters \( \beta \in \mathbb{R}^d \) and \( \alpha > 0 \). To do a maximum likelihood estimation, we have to maximize the complete data log-likelihood:

\[
\ell(D, \beta, \alpha) = \sum_{t \in D_0} \beta^T x_{j_t} + \sum_{t \in D_0} (\beta^T x_0 + \alpha \sum_{j \in S_t} e^{\beta^T x_j}) - \sum_{t=1}^T \log \left( e^{\beta^T x_0 e^{\alpha \sum_{j \in S_t} e^{\beta^T x_j}}} + \sum_{k \in S_t} e^{\beta^T x_k} \right),
\]

where \( D_0 \) is defined as the subset of \( D \) where there was no purchase.

**Lemma 4.5.** The log-likelihood function is not concave in \((\beta, \alpha)\) and hence the MLE problem is not a convex optimization problem.

In view of the above lemma, we prove the following statements and give an alternating projection maximization to get a local maximum.

**Lemma 4.6.** For a given value of \( \alpha \), the maximization problem over \( \beta \) can be reformulated as a convex optimization problem.

The partial maximization problem is:

\[
\max_{\beta} \sum_{t \in D_0} \beta^T x_{j_t} + \sum_{t \in D_0} (\beta^T x_0 + \alpha \sum_{j \in S_t} e^{\beta^T x_j}) - \sum_{t=1}^T \log \left( e^{\beta^T x_0 e^{\alpha \sum_{j \in S_t} e^{\beta^T x_j}}} + \sum_{k \in S_t} e^{\beta^T x_k} \right).
\]

We introduce the following new variables:

\[ z_t = \beta^T x_0 + \alpha \sum_{j \in S_t} e^{\beta^T x_j}, \quad t = 1, \ldots, T \]

Then we can re-write the above maximization as:

\[
\max_{\beta, z_t} \sum_{t \in D_0} \beta^T x_{j_t} + \sum_{t \in D_0} z_t - \sum_{t=1}^T \log \left( e^{z_t} + \sum_{k \in S_t} e^{\beta^T x_k} \right)
\]

s.t. \( \beta^T x_0 + \alpha \sum_{j \in S_t} e^{\beta^T x_j} - z_t = 0, \quad t = 1, \ldots, T \)

The objective function is now concave in \((\beta, z_t)\) as it is a sum of linear functions of \( \beta \) and \( z_t \) and the
negative of log-sum-exp function. Also, the equality constraints are convex functions in \((\beta, z_t)\). Hence we can solve this optimization problem efficiently. We also note that we are only introducing \(T\) new variables and constraints.

**Lemma 4.7.** For a given value of \(\beta\), the log-likelihood function is strictly concave in \(\alpha\) and hence it is unimodal. So the maximization problem over \(\alpha\) can be easily solved.

For a given value of \(\beta\), the partial maximization problem is given by:

\[
\max_{\alpha} \sum_{t \in D_0} \alpha \sum_{j \in S_t} v_j - \sum_{t=1}^T \log \left( v_0 e^{\alpha \sum_{j \in S_t} v_j + \sum_{k \in S_t} v_k} \right)
\]

Defining \(c_t := \sum_{k \in S_t} v_k\), we can re-write this as

\[
\max_{\alpha} \sum_{t \in D_0} \alpha (v_0 + c_t) - \sum_{t=1}^T \log \left( v_0 e^{\alpha (v_0 + c_t) + c_t} \right)
\]

A simple derivative calculation shows the above function is strictly concave in \(\alpha\) and hence the maximization problem is easy to solve.

In accordance with the above results, an iterative algorithm for maximizing the log-likelihood is to keep maximizing over \(\alpha\) and \(\beta\) alternatively until convergence. This will lead to a local maximum. Since this type of alternative maximization algorithm is dependent on the initial point, a good initial point could be \(\beta_{MNL,MLE}\) which is the maximum likelihood estimate for the MNL model (which can be easily found as the MLE for MNL model is a convex optimization problem). In experiments, we found the algorithm to converge after a few iterations only, as evident from the plot in Figure 3.

### 5 FPTAS for the Generalized Multinomial Logit model

In this section we present a fully polynomial time approximation scheme (FPTAS) for the assortment optimization problem under the Generalized Multinomial Logit model presented in Section 4. Let \(v\) (resp. \(V\)) be the minimum (resp. maximum) value of the transition probabilities. We can assume that \(v > 0\). For any given \(\epsilon > 0\), we use the following set of guesses for the denominator:

\[
V_\epsilon = \{v(1 + \epsilon)^l \mid l = 0, ..., L\},
\]

where \(L = O(\log(nV/v)/\epsilon)\). The number of guesses is polynomial in the number of products and \(1/\epsilon\). For given guess \(h \in V_\epsilon\), we try to find the best revenue possible with

\[
\sum_{j \in S} v_j \leq h,
\]
Figure 3: Estimate of $\alpha$ vs number of iterations for the estimation algorithm. The feature vector was 4 dimensional and we had 10 products, i.e., $d = 4, n = 10$. We can see convergence after a few iterations of the alternating projection algorithm.

using a dynamic program. We consider the following discretized values of $v_j$ in multiples of $\epsilon h/n$:

$$\forall j \in N \quad \bar{v}_j = \left\lceil \frac{v_j}{\epsilon h/n} \right\rceil.$$  

Let $I = \lceil n/\epsilon \rceil + n$. For each $(i, k) \in [I] \times [n]$, let $R(i, k)$ be the maximum revenue of any subset $S \subseteq \{1, \ldots, k\}$ such that

$$\sum_{j \in S} \bar{v}_j \leq i.$$  

We compute $R(i, k)$ using the following dynamic program

$$R(i, 1) = \begin{cases} 
 v_1 p_1 & \text{if } \bar{v}_1 \leq i \\
 0 & \text{if } i \geq 0 \\
 -\infty & \text{otherwise}
\end{cases}$$

$$R(i, k + 1) = \max\{v_{k+1} p_{k+1} + R(i - \bar{v}_{k+1}, k), R(i, k)\}.$$  

Let $S_h$ be the subset corresponding to $R(I, n)$, that is the assortment $S_h$ that maximizes the sum $\sum_{j \in S} v_j p_j$ such that (1) is verified (for $i = I$). We then construct a set of candidate assortments $S_h$ for all guesses $h$. And we return the best revenue that we get from all the candidates in the set. We write below the Algorithm that details how we get the best assortment for the FPTAS.

**Theorem 5.1.** Algorithm 1 returns an $(1 - \epsilon)$-optimal solution to our assortment problem, and its running time is $O\left(\frac{2^2}{\epsilon^2} \log(nV/v)\right)$.
Algorithm 1 FPTAS for the Generalized Multinomial Logit model

procedure FPTASGENMNL($\epsilon, v$)
  for $h \in V, \ell$
    Compute the discretized coefficients $\bar{v}_j = \left\lceil \frac{v_j \ell}{\epsilon h / n} \right\rceil$
    Compute $R(i,k)$ for all $(i,k) \in [I] \times [n]$ using the dynamic program above
    Let $S_h$ be the subset corresponding to $R(I,n)$
  end for
  Let $\mathcal{C} = \bigcup_{h \in V} S_h$
  return the set $S^* \in \mathcal{C}$ that has the best revenue
end procedure

Proof. Let $S^*$ be the optimal solution to the assortment optimization problem. There exists $l$ such that

$$v(1 + \epsilon)^{l-1} \leq \sum_{j \in S^*} v_j \leq v(1 + \epsilon)^l.$$

Let $h = v(1 + \epsilon)^l$. Then

$$\sum_{j \in S^*} \frac{v_j}{\epsilon h / n} \leq \frac{n}{\epsilon h} = \frac{n}{\epsilon},$$

and rounding up gives us

$$\sum_{j \in S^*} \bar{v}_j \leq \left\lceil \frac{n}{\epsilon} \right\rceil + n = I.$$

Thus $S^*$ belongs to the set of assortments such that (1) is verified for $I$. Let $S_h$ be the assortment corresponding to $R(I,n)$ for the guess $h$, that is the one that maximizes $\sum_{j \in S} v_j p_j$ subject to (1). Then since $S^*$ satisfies (1)

$$\sum_{j \in S_h} v_j p_j \geq \sum_{j \in S^*} v_j p_j.$$

Moreover,

$$\sum_{j \in S_h} v_j \leq \epsilon h / n \sum_{j \in S_h} \bar{v}_j \leq h(1 + \epsilon + \epsilon / n) \leq h(1 + 2\epsilon).$$

Since $x \mapsto \frac{1}{x + v_0 e^a(v_0 + x)}$ is a decreasing function and

$$R(S_h) = \frac{\sum_{j \in S_h} v_j p_j}{\sum_{k \in S_h} v_k + v_0 e^{a(v_0 + \sum_{j \in S_h} v_j)}} \geq \frac{\sum_{j \in S_h} v_j p_j}{v(1 + \epsilon)^l(1 + 2\epsilon) + v_0 e^{a(v_0 + v(1 + \epsilon)^l(1 + 2\epsilon))}}.$$

Let us first show that there exists $\beta > 0$ such that

$$v(1 + \epsilon)^{l-1} + v_0 e^{a(v_0 + v(1 + \epsilon)^l-1)} \geq (v(1 + \epsilon)^l(1 + 2\epsilon) + v_0 e^{a(v_0 + v(1 + \epsilon)^l(1 + 2\epsilon))}) \times (1 - \beta \epsilon). \quad (1)$$

23
Indeed,

\[
v(1 + \epsilon)^{l-1} \geq v(1 + \epsilon)^{l}(1 + 2\epsilon)(1 - \beta\epsilon)
\]

\[\iff 1 \geq (1 + \epsilon)(1 + 2\epsilon)(1 - \beta\epsilon) = 1 + (3 - \beta)\epsilon + (2 - 3\beta)\epsilon^2 - \beta\epsilon^3,
\]

which is clearly true at least for \(\beta \geq 3\). Moreover,

\[
v_0e^{\alpha(v_0+v(1+\epsilon)^{l-1})} \geq v_0e^{\alpha(v_0+v(1+\epsilon)(1+2\epsilon))} \times (1 - \beta\epsilon)
\]

\[\iff e^{\alpha(1+\epsilon)^{l-1}(1-(1+\epsilon)(1+2\epsilon))} \geq 1 - \beta\epsilon.
\]

Note that \(e^{\alpha(1+\epsilon)^{l-1}(1-(1+\epsilon)(1+2\epsilon))} = 1 - 3\alpha\epsilon + o(\epsilon)\). Therefore if we take \(\beta \geq \max(3, 3\alpha\epsilon)\), then the inequality (2) is verified. Consequently

\[
\frac{1}{v(1 + \epsilon)^{l}(1 + 2\epsilon) + v_0e^{\alpha(v_0+v(1+\epsilon)^{l-1})}} \geq \frac{1 - \beta\epsilon}{v(1 + \epsilon)^{l-1} + v_0e^{\alpha(v_0+v(1+\epsilon)^{l-1})}} \geq \frac{1 - \beta\epsilon}{\sum_{k \in S^*} v_k + v_0e^{\alpha(v_0+\sum_{k \in S^*} v_k)}}.
\]

by definition of \(l\). Therefore,

\[
R(S_h) \geq (1 - \beta\epsilon)\frac{\sum_{j \in S_h} v_j p_j}{\sum_{k \in S^*} v_k + v_0/e^{\alpha(v_0+\sum_{k \in S^*} v_k)}}
\]

\[
R(S_h) \geq (1 - \beta\epsilon)R(S^*),
\]

where in the last inequality we used that \(\sum_{j \in S_h} v_j p_j \geq \sum_{j \in S^*} v_j p_j\). This proves that our algorithm returns an \((1 - \epsilon)\)-optimal solution to the assortment problem.

**Running time** We try \(L = O((\log(nV/v)/\epsilon))\) guesses \(h\), and for each guess we formulate a dynamic programming with \(O(n^2/\epsilon)\) steps. Consequently the running time of the algorithm is \(O \left(\frac{n^2}{\epsilon^2} \log(nV/v)\right)\)

which is polynomial in the input size \(n\) and \(\frac{1}{\epsilon}\).

### 6 Generalized Markov Chain for Mixtures of Multinomial Logit models

In the previous model, we supposed that the initial transition matrix is of rank one. In this section, we will consider a more complex model, where the initial transition matrix is of low rank.

#### 6.1 Model’s hypothesis

**Generalized Markov chain model with Low Rank matrix** Let \(K < n\), we suppose that the initial transition matrix is of rank \(K\). There exists \((u^k)_{k \in [K]} \in ((0, 1)^n)^K\) and \((v^k)_{k \in [K]} \in ((0, 1)^{n+1})^K\) such that for all \(i \in [n]\), \(\sum_{j=0}^{n} \sum_{k=1}^{K} u^k_i v^k_j = 1\), and such that the initial transition matrix in the graph is given.
by
\[ \rho(N, N_+) = \left( \sum_{k \in [K]} u_k^i v_j^k \right)_{i \in [1, n], j \in [0, n]} = \sum_{k \in [K]} u^k (v^k)^T. \]

Given this initial transition probability matrix, we suppose that the probability of purchasing the product \( i \) while being at vertex \( i \) when \( S \) is offered is
\[
\mu^{LR}(i, S) = e^{-\alpha \left( \sum_{j \in S} \sum_{k \in [K]} u_k^i v_j^k \right)} = e^{-\alpha \left( \sum_{k \in [K]} u_k^i V_k(S) \right)},
\]
where \( V_k(S) = \sum_{j \in S} v_j^k \) for all \( k \in [K] \).

In this section, we will also make the following assumptions:

- we suppose that \( \forall l \in [n], \forall k \in [K] \) \( u_k^l v_k^l \leq \frac{1}{n} \) (by this, we mean that the probability of staying at the state \( l \) without buying the product \( l \) cannot be too high);

- we also suppose that \( \alpha \) is not too large compared to \( n \), more precisely, we suppose that \( \alpha \leq \log n \).

Under these assumptions, we can prove the following result:

**Lemma 6.1.** Let \( S \subseteq N \), we define the following matrix of \( \mathcal{M}_K(\mathbb{R}) \):
\[
UV(S) = \left( \sum_{l=1}^{n} (1 - \mu^{LR}(l, S)) u_l^i v_l^j \right)_{i,j \in [K]}.
\]

Then \( \rho(UV(S)) \leq 1 - \frac{1}{n} \).

**Proof.** Let \( ||.|| \) be the usual Euclidean norm, and \( \langle ., . \rangle \) its associated scalar product. We have that
\[
\rho(UV(S)) = \max_{x, ||x|| = 1} \langle UV(S)x, x \rangle.
\]

Let \( i \in [K] \),
\[
\langle UV(S)e_i, e_i \rangle = \sum_{l=1}^{n} (1 - \mu^{LR}(l, S)) u_l^i v_l^i \leq \frac{1}{n} \sum_{l=1}^{n} (1 - \mu^{LR}(l, S)),
\]
\[
\leq \frac{1}{n} \left( n - |S| + \sum_{l \in S} (1 - e^{-\alpha \sum_{k \in [K]} u_k^l V_k(S)}) \right),
\]
where the first inequality follows from the assumption made above on the coefficients \( u_k^l v_k^l \). Furthermore, since \( \alpha \leq \log n \),
\[
e^{-\alpha \sum_{k \in [K]} u_k^l V_k(S)} \geq e^{-\alpha \sum_{k \in [K]} u_k^l \sum_{j \in N_+} v_j^k} = e^{-\alpha} \geq \frac{1}{n}.
\]
Hence
\[
\langle UV(S)e_i, e_i \rangle \leq \frac{1}{n} \left( n - |S| + |S| \left( 1 - \frac{1}{n} \right) \right) = 1 - \frac{|S|}{n^2} \leq 1 - \frac{1}{n^2}.
\]
The inequality holds for all $e_i, i \in [n]$, and therefore $\forall x \in \mathbb{R}^K$ such that $||x|| = 1$. Hence,

$$\rho(UV(S)) \leq 1 - \frac{1}{n^2}$$

The fact that the spectral radius of $UV(S)$ is bounded away from 1 will be of use in the next subsection.

6.2 Assortment Optimization problem under the Generalized Markov chain model with low rank matrix

We compute the expected revenue that we get if we offer the subset $S$.

**Theorem 6.2.** Under the Generalized Markov chain with Low rank matrix model, the expected revenue that we get from offering the subset $S$ is

$$R^{LR}(S) = \sum_{i \in [n]} \lambda_i (1 - \mu^{LR}(i, S)) \sum_{k,k'=1}^K u^k_i (I - UV(S))^{-1}_{k'k} \left( \sum_{l \in S} \nu^l_{k'} \mu^{LR}(l, S) p_l \right)$$

$$+ \sum_{i \in S} \lambda_i \mu^{LR}(i, S) p_i$$

where $(I - UV(S))^{-1}_{k'k}$ is the coefficient of indices $k', k$ of the matrix $(I - UV(S))^{-1}$, where the matrix $UV(S) \in \mathcal{M}_K(\mathbb{R})$ is defined by

$$UV(S) = \left( \sum_{l=1}^n (1 - \mu^{LR}(l, S)) u^i_l v^j_l \right)_{i,j \in [K]}.$$

**Proof.** Let $S \in N$ be the chosen subset of products, and let $i \in S$. We recall from Section 3, that

$$\pi(i, S) = \lambda^T \left( I_n - \text{Diag}(1 - \mu(i, S)) \right) \rho(N, N)^{-1} \Pi(S)e_i.$$ (2)

This result still holds as the Generalized Multinomial Logit model with Low rank matrix has the same assumptions as the one needed to get this result. However, there is some difference with the assumptions made in Section 4. On the contrary of the Generalized Markov chain model presented in Section 4, we now have that

$$\mu^{LR}(i, S) = e^{-\alpha \times (\sum_{k \in [K]} u_i^k v^k(S))}$$

and

$$\rho(N, N) = \sum_{k \in [K]} u^k (v^k)^T.$$
where $v^k_s = (v^k_1, ..., v^k_n)$. Therefore we have to compute the coefficients of the matrix

$$\left(I_n - \text{Diag}((1 - \mu^{LR}(i, S)))\rho(N, N)\right)^{-1}$$

under this new model. We know that

$$\left(I_n - \text{Diag}((1 - \mu^{LR}(i, S)))\rho(N, N)\right)^{-1} = \sum_{l=0}^{\infty} \left(\text{Diag}((1 - \mu^{LR}(i, S)))\rho(N, N)\right)^l. \quad (3)$$

We will use the following notations:

- let $M := \text{Diag}((1 - \mu^{LR}(i, S)))\rho(N, N) = ( (1 - \mu^{LR}(i, S))\rho_{ij} )_{i,j \in [n]}$,
- let $UV(S) = \left( \sum_{l=1}^{n} (1 - \mu^{LR}(i, S))u^k_i v^k_j \right)_{i,j \in [K]}$,
- for $l \in \mathbb{N}$, we note $(UV(S)^l)_{k'k}$ the coefficient of indices $k', k$ of the matrix $UV(S)^l$, i.e. of the matrix $UV(S)$ elevated to the power of $l$.

First, we show by induction that

$$\forall l \geq 1 \quad M^l = \left( (1 - \mu^{LR}(i, S)) \sum_{k,k' \in [K]} u^k_i v^{k'}_j \left( UV(S)^{l-1} \right)_{k'k} \right)_{i,j \in [n]}.$$

**Initiation** For $l = 1$, we use the definition of $M$ and $\rho(N, N)$

$$M = \left( (1 - \mu^{LR}(i, S)) \sum_{k \in [K]} u^k_i v^k_j \right)_{i,j \in [n]}.$$

Since $UV(S)^0 = I_K$, we have that $(UV(S)^0)_{k'k} = \mathbb{1}_{k' = k}$. Therefore

$$M = \left( (1 - \mu^{LR}(i, S)) \sum_{k,k' \in [K]} u^k_i v^{k'}_j \left( UV(S)^0 \right)_{k'k} \right)_{i,j \in [n]}.$$

The result holds for $l = 1$.

**Inductive step** Suppose that the result holds for a certain $l \in \mathbb{N}^*$. We compute the coefficient of indices $i, j$ of $M^{l+1}$:

$$\left( M^{l+1} \right)_{ij} = \sum_{q=1}^{n} \left( M^l \right)_{iq} (1 - \mu^{LR}(q, S))\rho_{qj}$$
Using the induction hypothesis,

\[
\left( M^{l+1} \right)_{ij} = \sum_{q=1}^{n} \left( 1 - \mu^{LR}(i, S) \right) \sum_{k, k' \in [K]} u_{ik}^{q} v_{kj}^{k'} \left( UV(S)^{l-1} \right)_{k'k} \left( 1 - \mu^{LR}(q, S) \right) \sum_{k'' \in [K]} u_{qk''}^{k''} v_{kj}^{k''}
\]

\[
= (1 - \mu^{LR}(i, S)) \sum_{k, k' \in [K]} u_{ik}^{k} v_{kj}^{k'} \left( \sum_{q \in [n]} \left( 1 - \mu^{LR}(q, S) \right) u_{qk}^{k''} v_{kj}^{k''} \right) \left( UV(S)^{l-1} \right)_{k'k}
\]

\[
= (1 - \mu^{LR}(i, S)) \sum_{k, k' \in [K]} u_{ik}^{k} v_{kj}^{k'} \left( UV(S) \right)_{k'k}.
\]

Therefore the result holds for \( l + 1 \).

**Conclusion** For all \( l \geq 1 \), we have that

\[
M^l = \left( 1 - \mu^{LR}(i, S) \right) \sum_{k, k' \in [K]} u_{ik}^{k} v_{kj}^{k'} \left( UV(S)^{l-1} \right)_{k'k}
\]

Now that we have this result, we can compute the coefficient of indices \( i, j \) of the matrix \( (I_n - \text{Diag}((1 - \mu^{LR}(i, S)))\rho(N, N))^{-1} = (I_n - M)^{-1} \) using (2):

\[
((I_n - \text{Diag}((1 - \mu^{LR}(i, S)))\rho(N, N))^{-1})_{ij} = \sum_{l=0}^{\infty} (M^l)_{ij}
\]

\[
= 1_{i=j} + \sum_{l=1}^{\infty} \left( 1 - \mu^{LR}(i, S) \right) \sum_{k, k' \in [K]} u_{ik}^{k} v_{kj}^{k'} \left( UV(S)^{l-1} \right)_{k'k}
\]

\[
= 1_{i=j} + \left( 1 - \mu^{LR}(i, S) \right) \sum_{k, k' \in [K]} u_{ik}^{k} v_{kj}^{k'} \sum_{l=1}^{\infty} \left( UV(S)^{l-1} \right)_{k'k}
\]

\[
= 1_{i=j} + \left( 1 - \mu^{LR}(i, S) \right) \sum_{k, k' \in [K]} u_{ik}^{k} v_{kj}^{k'} \left( I_K - UV(S) \right)_{k'k}^{-1}
\]

. The last inequality holds since \( \rho(UV(S)) < 1 \), as shown in Lemma 6.1. Injecting it in (1), and
computing $\sum_{i \in S} \pi^{LR}(i, S) p_i$, we have

$$
R^{LR}(S) = \sum_{i=1}^{n} \lambda_i \sum_{j=1}^{n} \left( 1_{i=j} + (1 - \mu^{LR}(i, S)) \sum_{k,k' \in [K]} u_i^k v_j^{k'} ((I_K - UV(S))^{-1})_{k,k'} \right) \mu^{LR}(j, S) p_j
$$

$$
= \sum_{i=1}^{n} \lambda_i \mu^{LR}(i, S) p_i + \sum_{i=1}^{n} \lambda_i (1 - \mu^{LR}(i, S)) \sum_{j=1}^{n} \left( \sum_{k,k' \in [K]} u_i^k v_j^{k'} ((I_K - UV(S))^{-1})_{k,k'} \right) \mu^{LR}(j, S) p_j
$$

$$
= \sum_{i=1}^{n} \lambda_i \mu^{LR}(i, S) p_i + \sum_{i=1}^{n} \lambda_i (1 - \mu^{LR}(i, S)) \sum_{k,k' \in [K]} u_i^k ((I_K - UV(S))^{-1})_{k,k'} \sum_{j=1}^{n} v_j^{k'} \mu^{LR}(j, S) p_j
$$

$$
+ \sum_{i \in S} \lambda_i \mu^{LR}(i, S) \left( p_i - \sum_{k,k' \in [K]} u_i^k ((I_K - UV(S))^{-1})_{k,k'} \left( \sum_{j=1}^{n} v_j^{k'} \mu^{LR}(j, S) p_j \right) \right).
$$

Finally, since $\mu^{LR}(j, S) = 0, \forall j \notin S$, after reorganizing the terms, we can rewrite this as

$$
R^{LR}(S) = \sum_{i \in [n]} \lambda_i (1 - \mu^{LR}(i, S)) \sum_{k,k' = 1}^{K} u_i^k (I - UV(S))^{-1} \left( \sum_{l \in S} v_l^{k'} \mu^{LR}(l, S) p_l \right)
$$

$$
+ \sum_{i \in S} \lambda_i \mu^{LR}(i, S) p_i,
$$

which concludes the proof.

In view of this theorem, if we define

$$
f(i, S) := \sum_{k,k' = 1}^{K} u_i^k (I - UV(S))^{-1} \left( \sum_{l \in S} v_l^{k'} \mu^{LR}(l, S) p_l \right),
$$

then the assortment optimization problem under the Generalized Markov chain model with Low rank matrix is given by

$$
\max_{S \subseteq N} R^{LR}(S) := \max_{S \subseteq N} \sum_{i \in [n]} \lambda_i (1 - \mu^{LR}(i, S)) f(i, S) + \sum_{i \in S} \lambda_i \mu^{LR}(i, S) p_i.
$$

### 6.3 FPTAS for the Generalized Markov chain model with low rank matrix

We will present in this subsection a FPTAS for the Generalized Markov chain model with low rank matrix. To prove that the FPTAS provides a $(1 - \epsilon)$-optimal solution to our assortment optimization problem, we will first need to prove that the spectral radius of $UV(S)$ is bounded away from 1. Let us first show this result in the following lemma.

**Lemma 6.3.** Let $S \subseteq N$. Suppose that $\exists H = (h_{kk'})_{k,k' \in [K]} \in \mathcal{M}_K(\mathbb{R})$ and $\exists \delta, \delta' > 0$ such that
\[ h_{kk'}(1 - \delta') \leq UV(S)_{kk'} \leq h_{kk'}(1 + \delta). \]

Then we have that
\[
((I_K - H)^{-1})_{kk'} (1 - \delta' X(h)) \leq ((I_K - UV(S))^{-1})_{kk'} \leq ((I_K - H)^{-1})_{kk'} (1 + \delta X(H)),
\]

where \( X \) is a function of the elements of \( H \).

**Proof.** From Lemma 6.1, we know that \( \rho(UV(S)) \leq 1 - \frac{1}{n^2} < 1 \). Therefore we can write for all \( k, k' \in [K] \)

\[
((I_K - UV(S))^{-1})_{kk'} = \sum_{l=0}^{\infty} \left( UV(S)^l \right)_{kk'},
\]

We show by induction that \( \forall l \in \mathbb{N}, \forall k, k' \in [K], (UV(S)^l)_{kk'} \leq (H^l)_{kk'} (1 + \delta)^l \). For \( l = 0 \), the inequality is trivial. Suppose that the inequality holds for all \( m \leq l \) for a certain \( l \in \mathbb{N} \). Then

\[
\left( UV(S)^{l+1} \right)_{kk'} = \sum_{s=1}^{K} \left( UV(S)^l \right)_{ks} (UV(S))_{sk'} \\
\leq \sum_{s=1}^{K} \left( H^l \right)_{ks} (1 + \delta)^l H_{sk'} (1 + \delta) = \left( H^{l+1} \right)_{kk'} (1 + \delta)^{l+1},
\]

which concludes the proof. Therefore,

\[
((I_K - UV(S))^{-1})_{kk'} \leq \sum_{l=0}^{\infty} \left( H^l \right)_{kk'} (1 + \delta)^l = ((I_K - (1 + \delta)H)^{-1})_{kk'}.
\]

Likewise, we can easily show that for all \( k, k' \in [K] \)

\[
((I_K - UV(S))^{-1})_{kk'} \geq ((I_K - (1 - \delta')H)^{-1})_{kk'}.
\]

Now, using the differential of the inverse of a matrix, we have

\[
(I_K - (1 + \delta)H)^{-1} = (I_K - H)^{-1} + (I_K - H)^{-1} \delta H (I_K - H)^{-1} + (I_K - H)^{-2} \delta^2 H^2 (I_K - H)^{-1} \epsilon(\delta H),
\]

where \( \epsilon(\delta) \to 0 \). Therefore

\[
((I_K - (1 + \delta)H)^{-1})_{kk'} = ((I_K - H)^{-1})_{kk'} + \delta \left( (I_K - H)^{-1} H (I_K - H)^{-1} \right)_{kk'} + o(\delta).
\]

We finally deduce

\[
((I_K - UV(S))^{-1})_{kk'} \leq ((I_K - (1 + \delta)H)^{-1})_{kk'} \leq ((I_K - H)^{-1})_{kk'} (1 + \delta X(H)),
\]
where \( X(H) := 1 + \max_{k,k'} \left\{ \frac{(I_K-H)^{-1}H(I_K-H)^{-1})_{kk'}}{(I_K-H)^{-1})_{kk'}} \right\} \). We can prove the lower bound using the same argument.

We will now present the FPTAS. Let \( v^k \) (resp. \( V^k \)) be the minimum (resp. maximum) value of \( \{v^k_i\}_{i \in N} \) for all \( k \in [K] \). We can assume that \( v^k > 0 \). For any given \( \epsilon > 0 \), we use the following sets of guesses:

\[
W^k_\epsilon = \{v^k(1 + \epsilon)^t, t = 0, ..., T^k\}, \text{ for all } k \in [K],
\]

where \( T^k = O(\log(nV^k/v^k)/\epsilon) \). A guess \( h \) belongs in the set

\[
W_\epsilon = W^1_\epsilon \times ... \times W^K_\epsilon.
\]

The number of guesses is polynomial in the input size and \( 1/\epsilon \). For given guess \( h = (h_1, ..., h_K) \in W_\epsilon \), we try to find the best revenue possible with

\[
h_k \leq \sum_{j \in S} v^k_j \leq h_k(1 + \epsilon), \text{ for all } k \in [K],
\]

using a dynamic program. We consider the following discretized values of \( v^k_j \) in multiples of \( \epsilon h_k/n \):

\[
\forall k \in [K], \forall j \in N, \quad \bar{v}^k_j = \left\lfloor \frac{v^k_j}{\epsilon h_k/n} \right\rfloor.
\]

We will denote \( \bar{v}_j \) the vector \( \bar{v}_j := (\bar{v}^1_j, ..., \bar{v}^K_j) \). Let \( L = [n/\epsilon] \), and \( U = [n/\epsilon] + n \). The dynamic program will try to maximize the total expected revenue that we will denote as \( R^{DP} \). For each \( (l, u, m) \in [L]^K \times [U]^K \times [n] \), let \( R^{DP}(l, u, m) \) be the maximum revenue \( R^{DP} \) of any subset \( S \subseteq \{1, ..., m\} \) such that

\[
l_k \leq \sum_{j \in S} \bar{v}^k_j \leq u_k \quad \forall k \in [K]
\]

We define, for each guess \( h \),

\[
\mu_i(h) := e^{-\alpha(\sum_{k=1}^K h_k u^k_i)} \quad \forall i \in N,
\]

For each guess \( h \in W_\epsilon \), \( \mu(h) = (\mu_1(h), ..., \mu_n(h)) \) will therefore be an estimation of the value of the \( \mu^{LR}(i, S) \)'s. Given this, we will also use the following estimation \( H \) of the matrix \( UV(S) \) defined before:

\[
H(h) := \left( \sum_{i=1}^n (1 - \mu_i(h)) \bar{v}^k_i \bar{v}^{k'}_i \right)_{k,k'[\in[K]}.
\]
Finally, let us also consider for each $i \in N$ and each $S$, the following estimate for $f(i, S)$ defined in (4)

$$f_i(h, S) = \sum_{k, k' = 1}^{K} u_i^k (I - H(h))^{-1}_{k,k'} \left( \sum_{j \in S} v_j^{k'} \mu_j(h)p_j \right).$$

For each guess $h$, we define the approximate revenue of the subset $S$ as

$$R_{DP}(h, S) = \sum_{i \in S} \lambda_i \mu_i(h)p_i + \sum_{i = 1}^{n} \lambda_i (1 - \mu_i(h))f_i(h, S).$$

We compute $R_{DP}(l, u, m)$ using the following dynamic program

$$R_{DP}(l, u, 1) = \begin{cases} \lambda_1 \mu_1(h)p_1 + \sum_{i = 1}^{n} \lambda_i (1 - \mu_i(h)) \sum_{k, k' = 1}^{K} u_i^k (I - H(h))^{-1}_{k,k'} v_1^{k'} \mu_1(h)p_1 & \text{if } l \leq \bar{v}_1 \leq u \\ 0 & \text{if } l \leq 0 \text{ and } u \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$R_{DP}(l, u, m) = \max \left\{ \lambda_m \mu_m(h)p_m + \sum_{i = 1}^{n} \lambda_i (1 - \mu_i(h)) \sum_{k, k' = 1}^{K} u_i^k (I - H(h))^{-1}_{k,k'} v_m^{k'} \mu_m(h)p_m \\ + R_{DP}(l - \bar{v}_m, u - \bar{v}_m, m - 1), R_{DP}(l, u, m - 1) \right\}.$$ 

We note that the number of states in the above dynamic program is $O \left( \left( \frac{u}{\epsilon} \right)^{2K} n \right)$ and each step of the dynamic program does require a summation but that can still be done in $O(nK^2)$ time, which results in a total run time of $O \left( \left( \frac{u}{\epsilon} \right)^{2K} n^2 K^2 \right).$

Let $S_h$ be the subset corresponding to $R_{DP}(L, U, n)$. We then construct a set of candidate assortments $S_h$ for all guesses $h$. And we return the best revenue that we get from all the candidates in the set.

We write below the Algorithm that details how we get the best assortment for the FPTAS.

**Algorithm 2 FPTAS for the Generalized Markov chain model with Low rank matrix**

```
procedure FPTASGenMxtMNL(\epsilon, u^k, v^k) 
  for h \in W_{\epsilon} do 
    Compute the discretized coefficients \( \bar{v}_j^k = \left\lceil \frac{v_j^k}{\epsilon h_k/n} \right\rceil \)
    Compute $R_{DP}(l, u, m)$ for all $(l, u, m) \in [L]^K \times [U]^K \times [n]$ using the dynamic program above
    Let $S_h$ be the subset corresponding to $R_{DP}(L, U, n)$
  end for 
  Let $C = \cup_{h \in W_{\epsilon}} S_h$
  return the set $S^* \in C$ that has the best revenue
end procedure
```
Theorem 6.4. Algorithm 2 returns an \((1 - \epsilon)\)-optimal solution to our assortment problem, and its running time is \(O \left( \frac{2^{2K^2} + 2}{e^{\alpha}}K^2 \log(nV/v)^K \right)\).

Proof. Let \(S^*\) be the optimal solution to the assortment optimization problem. There exist \(t_1, \ldots, t_K\) such that for all \(k \in [K]\)
\[
v^k(1 + \epsilon)^t_k \leq \sum_{j \in S^*} v^k_j =: V^k(S^*) \leq v^k(1 + \epsilon)^{t_k + 1}.
\]

Let \(h = (v^1(1 + \epsilon)^{t_1}, \ldots, v^K(1 + \epsilon)^{t_K})\). Choose the set \(S_h\) that maximizes the dynamic program defined above. Then by definition of \(\bar{v}^k_j\), for each \(k \in [K]\)
\[
V^k(S_h) := \sum_{j \in S_h} v^k_j \leq \epsilon h_k/n \sum_{j \in S_h} \bar{v}^k_j \leq \frac{\epsilon h_k}{n} U = \frac{\epsilon h_k}{n} ([n/\epsilon] + n) \leq h_k(1 + 2\epsilon),
\]
and
\[
V^k(S_h) := \sum_{j \in S_h} v^k_j \geq \epsilon h_k/n \sum_{j \in S_h} (\bar{v}^k_j - 1) \geq \frac{\epsilon h_k}{n} (L - |S_h|) \geq \frac{\epsilon h_k}{n} ([n/\epsilon] - n) \geq h_k(1 - \epsilon).
\]

First of all, since for all \(k \in [K]\), \(h_k \leq V^k(S^*) \leq h_k(1 + \epsilon),\)
\[
e^{-\alpha \sum_{k=1}^K h_k u^k_j (1 + \epsilon)} \leq \mu^L R(j, S^*) = e^{-\alpha \sum_{k=1}^K V^k(S^*) u^k_j} \leq e^{-\alpha \sum_{k=1}^K h_k u^k_j} = \mu_j(h).
\]

Thus, if we note \(\sum_{k=1}^K h_k u^k_j =: \langle h, u_j \rangle\),
\[
\mu_j(h)(1 - \epsilon \alpha \langle h, u_j \rangle) \leq \mu^L R(j, S^*) \leq \mu_j(h), \tag{5}
\]
and using the same arguments for \(h_k(1 - \epsilon) \leq V^k(S_h) \leq h_k(1 + 2\epsilon)\), we can show that
\[
\mu_j(h)(1 - 2\epsilon \alpha \langle h, u_j \rangle) \leq \mu^L R(j, S_h) \leq \mu_j(h)(1 + \epsilon \alpha \langle h, u_j \rangle).
\]

Therefore
\[
UV(S^*)_{kk'} = \sum_{i \in [n]} (1 - \mu^L R(i, S^*)) u^k_i v^{k'}_i \leq \sum_{i \in [n]} (1 - \mu_i(h)) u^k_i v^{k'}_i \left(1 + \frac{\mu_i(h)}{1 - \mu_i(h)} \epsilon \alpha \langle h, u_i \rangle \right)
\]
\[
\leq \sum_{i \in [n]} (1 - \mu_i(h)) u^k_i v^{k'}_i \left(1 + \epsilon \alpha \max_{i \in [n]} \left\{ \frac{\mu_i(h)}{1 - \mu_i(h)} \langle h, u_i \rangle \right\} \right)
\]

Hence we get for all \(k, k' \in [K]\)
\[
H(h)_{kk'} \leq UV(S^*)_{kk'} \leq H(h)_{kk'}(1 + \delta), \tag{6}
\]

33
We want to prove that $R$ and $Recall the expression of optimal revenue $f$. Let us now compare $f$ using bounds from (5) and (7). We have $\delta$, where $k, k' \in [K]$

$$\left((I - H(h))^{-1}\right)_{kk'} \leq \left((I - UV(S^*))^{-1}\right)_{kk'} \leq \left((I - H(h))^{-1}\right)_{kk'} (1 + \delta X(H)), \quad (7)$$

and

$$\left((I - H(h))^{-1}\right)_{kk'} (1 - \delta X(h)) \leq \left((I - UV(S^*))^{-1}\right)_{kk'} \leq \left((I - H(h))^{-1}\right)_{kk'} (1 + 2\delta X(H)).$$

We want to prove that $R^{LR}(S_h) \geq (1 - g(\epsilon))R^{LR}(S^*)$ for a certain function $g$ such that $g(x) \to 0$.

Recall the expression of optimal revenue

$$R^{LR}(S^*) = \sum_{i \in S^*} \lambda_i \mu^{LR}(i, S^*) p_i + \sum_{i=1}^n \lambda_i (1 - \mu^{LR}(i, S^*)) f(i, S^*).$$

Let us now compare $f_i(h, S^*)$ and $f(i, S^*) := \sum_{k,k'=1}^K u_i^k (I - UV(S^*))^{-1}_{kk'} \left(\sum_{j=1}^n v_{ij}^{k'} \mu^{LR}(j, S^*) p_j\right)$ using bounds from (5) and (7). We have

$$f_i(h, S^*)(1 - \delta_2) \leq f(i, S^*) := \sum_{k,k'=1}^K u_i^k (I - UV(S^*))^{-1}_{kk'} \left(\sum_{j=1}^n v_{ij}^{k'} \mu^{LR}(j, S^*) p_j\right) \leq f_i(h, S^*)(1 + \delta X(H)),$$

where $\delta_2 := \varepsilon \alpha \max_{i \in [n]} \{\langle h, u_i \rangle\}$. Using the bounds for $\mu^{LR}(j, S_h)$ and $f(i, S_h)$, we also have

$$f_i(h, S_h)(1 - \delta X(h))(1 - 2\delta_2) \leq f(i, S_h) \leq f_i(h, S_h)(1 + 2\delta X(H))(1 + \delta_2).$$

Using these upper and lower bounds for $f(i, S^*)$ (resp. $f(i, S_h)$) and the corresponding lower and upper bounds for $\mu^{LR}(j, S^*)$ (resp. $\mu^{LR}(j, S_h)$) from equation (5), we have

$$R^{LR}(S^*) \leq \sum_{i \in S^*} \lambda_i \mu_i(h) p_i + \sum_{i=1}^n \lambda_i (1 - \mu_i(h))(1 + \delta) f_i(h, S^*) (1 + \delta X(H)) \leq (1 + \delta) (1 + \delta X(H)) R^{DP}(h, S^*),$$

34
and
\[ R^{LR}(S_h) \geq \sum_{i \in S^*} \lambda_i (1 - 2\delta_2) \mu_i(h) p_i + \sum_{i=1}^{n} \lambda_i (1 - \mu_i(h))(1 - \delta)(1 - \delta X(h))(1 - 2\delta_2) f_i(h, S_h) \]
\[ \geq (1 - \delta)(1 - \delta X(h))(1 - 2\delta_2) R^{DP}(h, S_h). \]

Next, by the definition of \( S_h \) being the assortment that maximizes the dynamic program for the given guess \( h \), we also have \( R^{DP}(h, S^*) \leq R^{DP}(h, S_h) \). Combining this with the bounds derived above, we get
\[ R^{LR}(S^*) \leq \frac{(1 + \delta)(1 + \delta X(H))}{(1 - \delta)(1 - \delta X(h))(1 - 2\delta_2)} R^{LR}(S_h). \] (8)

Hence we get the following lower bound for the expected revenue obtained by the assortment \( S_h \):
\[ \frac{(1 - \delta)(1 - \delta X(h))(1 - 2\delta_2)}{(1 + \delta)(1 + \delta X(H))} R^{LR}(S^*) \leq R^{LR}(S_h). \]

Finally, to argue that this gives us \((1 - \beta \epsilon)\) solution, since \( \delta X(H) > 0 \), to conclude the proof, we need to show that \( \frac{(1 - \delta)(1 - \delta X(h))(1 - 2\delta_2)}{(1 + \delta)(1 + \delta X(H))} \) is not too far from 1. This holds since,
\[ \delta := C_1(h) \epsilon \quad \text{and} \quad \delta_2 := C_2(h) \epsilon. \]

Thus
\[ \delta X(H) = C_3(h) \epsilon, \]
and
\[ \frac{(1 - \delta)(1 - \delta X(h))(1 - 2\delta_2)}{(1 + \delta)(1 + \delta X(H))} = 1 - 2(C_1(h) + C_2(h) + C_3(h)) \epsilon + o(\epsilon) \geq 1 - \beta \epsilon, \]
where \( \beta \) is a constant and \( C_1(h), C_2(h), C_3(h) \) are polynomial functions of the \( h_k \)’s. The last inequality holds because \( C_1(h), C_2(h) \) and \( C_3(h) \) can be easily bounded over \( W_\epsilon \). The bound may depend on \( \epsilon \), but since we multiply it by \( \epsilon \), it will again be an \( o(\epsilon) \) term, and is thus of no importance.

We have finally proven that our algorithm returns an assortment \( S_h \) that is a \((1 - \epsilon)\)-optimal solution to our assortment problem.
\[ R^{LR}(S^*)(1 - \beta \epsilon) \leq R^{LR}(S_h) \leq R^{LR}(S^*) \] (9)

**Running time** We try a total of \( \prod_{k \in [K]} O(\log(nV^k/v^k)/\epsilon) = O((\log(nV/v)/\epsilon)^K) \) guesses for \( h \) (where \( V := \max_{k \in [K]} V^k \) and \( v := \min_{k \in [K]} v^k \)). For each guess we formulate a dynamic programming with \( O \left( \left( \frac{n}{\epsilon} \right)^{2K} n^2 K^2 \right) \) run-time. Consequently the running time of the algorithm is \( O \left( \frac{n^{2(K+1)} K^2}{\epsilon n} \log^K (nV/v) \right) \) which is polynomial in the input size \( n \) and \( \frac{1}{\epsilon} \). \\[ \square \]
7 Conclusion

Our main contribution in this paper is to build upon the Markovian chain based model presented by Blanchet et al. [2] and present a generalized model that addresses two significant limitations of existing random utility and rank-based choice models in capturing dynamic preferences and the choice overload phenomenon.

The Generalized Markov chain model attempts to capture both dynamic preferences and choice overload phenomenon by considering a modified choice or selection process, where a customer stops at a state corresponding to an offered product with some probability that depends on the set of offered products. This implicitly models the search cost in the selection process and therefore, captures both dynamic preferences and the choice overload phenomenon. Therefore, we present a novel framework to overcome the limitations in the existing choice models.

We also consider the special case of rank-1 Markov chain, referred to as the Generalized MNL model. We show that the optimization problem under this model is NP-hard and present a fully polynomial time approximation scheme for this problem. We also present a fully polynomial time approximation scheme for this problem when we suppose that the transition matrix is of low rank. The first model generalizes the MNL model while the second model generalizes the Mixture of MNL model. An important future step would be to find a set of data, when first a set $S$ is offered to the customers and then another set $S'$ is offered to the same customers, to verify that our model captures better customers’ choice than the MNL or the Mixture of MNL models. For the moment, our model has shown to be reliable for random data.

8 Bibliography

References

[1] Moshe E Ben-Akiva and Steven R Lerman. *Discrete choice analysis: theory and application to travel demand*, volume 9. MIT press, 1985.

[2] Jose Blanchet, Guillermo Gallego, and Vineet Goyal. A markov chain approximation to choice modeling. *Operations Research*, 64(4):886–905, 2016.

[3] Hakjin Chung, Hyun-Soo Ahn, and Stefanus Jasim. On (re-scaled) multi-attempt approximation of customer choice model and its application to assortment optimization. 2016.

[4] James Davis, Guillermo Gallego, and Huseyin Topaloglu. Assortment planning under the multinomial logit model with totally unimodular constraint structures. *Department of IEOR, Columbia University. Available at [http://www.columbia.edu/~gmg2/logit_const.pdf](http://www.columbia.edu/~gmg2/logit_const.pdf)*, 2013.

[5] James M Davis, Guillermo Gallego, and Huseyin Topaloglu. Assortment optimization under variants of the nested logit model. *Operations Research*, 62(2):250–273, 2014.

[6] Antoine Desir, Vineet Goyal, Srikanth Jagabathula, and Danny Segev. Assortment optimization under the mallows model. In *Advances in Neural Information Processing Systems*, pages 4700–4708, 2016.

[7] Antoine Désir, Vineet Goyal, and Jiawei Zhang. Capacitated assortment optimization: Hardness and approximation. 2017.

[8] Vivek F Farias, Srikanth Jagabathula, and Devavrat Shah. A nonparametric approach to modeling choice with limited data. *Management Science*, 59(2):305–322, 2013.
[9] Guillermo Gallego, Garud Iyengar, R Phillips, and Abha Dubey. Managing flexible products on a network. 2004.
[10] Sheena S Iyengar and Mark R Lepper. When choice is demotivating: Can one desire too much of a good thing? Journal of personality and social psychology, 79(6):995, 2000.
[11] S Jagabathula and P Rusmevichientong. A two-stage model of consideration set and choice: Learning, revenue prediction. Technical report, and applications. Working paper, 2013.
[12] Srikanth Jagabathula and Gustavo Vulcano. A partial-order-based model to estimate individual preferences using panel data. Management Science, 2017.
[13] R Duncan Luce. Individual choice behavior: A theoretical analysis. Courier Corporation, 2005.
[14] Daniel McFadden. Modeling the choice of residential location. Transportation Research Record, (673), 1978.
[15] Kameng Nip, Zhenbo Wang, and Zizhuo Wang. Assortment optimization under a single transition model. 2017.
[16] Robin L Plackett. The analysis of permutations. Applied Statistics, pages 193–202, 1975.
[17] Paat Rusmevichientong, David Shmoys, Chaoxu Tong, and Huseyin Topaloglu. Assortment optimization under the multinomial logit model with random choice parameters. Production and Operations Management, 2014.
[18] Benjamin Scheibehenne, Rainer Greifeneder, and Peter M Todd. Can there ever be too many options? a meta-analytic review of choice overload. Journal of Consumer Research, 37(3):409–425, 2010.
[19] Kalyan Talluri and Garrett Van Ryzin. Revenue management under a general discrete choice model of consumer behavior. Management Science, 50(1):15–33, 2004.
[20] Garrett van Ryzin and Gustavo Vulcano. An expectation-maximization method to estimate a rank-based choice model of demand. Operations Research, 65(2):396–407, 2017.
[21] Ruxian Wang and Ozge Sahin. The impact of consumer search cost on assortment planning and pricing. 2014.
[22] Ruxian Wang and Zizhuo Wang. Consumer choice models with endogenous network effects. Management Science, 2016.
[23] Huw CWL Williams. On the formation of travel demand models and economic evaluation measures of user benefit. Environment and planning A, 9(3):285–344, 1977.
[24] Dan Zhang and William L Cooper. Revenue management for parallel flights with customer-choice behavior. Operations Research, 53(3):415–431, 2005.
This figure "ProbaBuying.png" is available in "png" format from:

http://arxiv.org/ps/1911.06716v2
This figure "Signature.png" is available in "png" format from:

http://arxiv.org/ps/1911.06716v2
This figure "complete1.png" is available in "png" format from:

http://arxiv.org/ps/1911.06716v2
This figure "complete2.png" is available in "png" format from:

http://arxiv.org/ps/1911.06716v2