

RESEARCH ARTICLE

A regret lower bound for assortment optimization under the capacitated MNL model with arbitrary revenue parameters

Yannik Peeters¹  and Arnoud V. den Boer^{1,2}

¹ Amsterdam Business School, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands.
E-mail: y.a.peeters@uva.nl

² Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 105-107, 1098 XG Amsterdam, The Netherlands.

Keywords: Assortment optimization, Incomplete information, Multinomial logit model, Regret lower bound

Abstract

In this note, we consider dynamic assortment optimization with incomplete information under the capacitated multinomial logit choice model. Recently, it has been shown that the regret (the cumulative expected revenue loss caused by offering suboptimal assortments) that any decision policy endures is bounded from below by a constant times \sqrt{NT} , where N denotes the number of products and T denotes the time horizon. This result is shown under the assumption that the product revenues are *constant*, and thus leaves the question open whether a lower regret rate can be achieved for nonconstant revenue parameters. In this note, we show that this is not the case: we show that, for *any* vector of product revenues there is a positive constant such that the regret of any policy is bounded from below by this constant times \sqrt{NT} . Our result implies that policies that achieve $O(\sqrt{NT})$ regret are asymptotically optimal for all product revenue parameters.

1. Introduction

In this note, we consider the problem of assortment optimization under the multinomial logit (MNL) choice model with a capacity constraint on the size of the assortment and incomplete information about the model parameters. This problem has recently received considerable attention in the literature (see, e.g., [1–7,9,10]).

Two notable recent contributions are from Agrawal *et al.* [1,2], who construct decision policies based on Thompson Sampling and Upper Confidence Bounds, respectively, and show that the regret of these policies—the cumulative expected revenue loss compared with the benchmark of always offering an optimal assortment—is, up to logarithmic terms, bounded by a constant times \sqrt{NT} , where N denotes the number of products and $T \geq N$ denotes the length of the time horizon. These upper bounds are complemented by the recent work from Chen and Wang [3], who show that the regret of any policy is bounded from below by a positive constant times \sqrt{NT} , implying that the policies by Agrawal *et al.* [1,2] are (up to logarithmic terms) asymptotically optimal.

The lower bound by Chen and Wang [3] is proven under the assumption that the product revenues are *constant*—that is, each product generates the same amount of revenue when sold. In practice, it often happens that different products have *different* marginal revenues, and it is *a priori* not completely clear whether the policies by Agrawal *et al.* [1,2] are still asymptotically optimal or that a lower regret can be achieved. In addition, Chen and Wang [3] assume that K , the maximum number of products allowed in an assortment, is bounded by $\frac{1}{4} \cdot N$, but point out that this constant $\frac{1}{4}$ can probably be increased.

In this note, we settle this open question by proving a \sqrt{NT} regret lower bound for *any* given vector of product revenues. This implies that policies with $O(\sqrt{NT})$ regret are asymptotically optimal *regardless*

of the product revenue parameters. Furthermore, our result is valid for all $K < \frac{1}{2}N$, thereby confirming the intuition of Chen and Wang [3] that the constraint $K \leq \frac{1}{4}N$ is not tight.

2. Model and main result

We consider the problem of dynamic assortment optimization under the MNL choice model. In this model, the number of products is $N \in \mathbb{N}$. Henceforth, we abbreviate the set of products $\{1 \dots, N\}$ as $[N]$. Each product $i \in [N]$ yields a known marginal revenue to the seller of $w_i > 0$. Without loss of generality due to scaling, we can assume that $w_i \leq 1$ for all $i \in [N]$. Each product $i \in [N]$ is associated with a preference parameter $v_i \geq 0$, unknown to the seller. Each offered assortment $S \subseteq [N]$ must satisfy a capacity constraint, that is, $|S| \leq K$ for capacity constraint $K \in \mathbb{N}$, $K \leq N$. For notational convenience, we write

$$\mathcal{A}_K := \{S \subseteq [N] : |S| \leq K\}$$

for the collection of all assortments of size at most K , and

$$\mathcal{S}_K := \{S \subseteq [N] : |S| = K\}$$

for the collection of all assortments of exact size K . Let $T \in \mathbb{N}$ denote a finite time horizon. Then, at each time $t \in [T]$, the seller selects an assortment $S_t \in \mathcal{A}_K$ based on the purchase information available up to and including time $t - 1$. Thereafter, the seller observes a single purchase $Y_t \in S_t \cup \{0\}$, where product 0 indicates a no-purchase. The purchase probabilities within the MNL model are given by

$$\mathbb{P}(Y_t = i | S_t = S) = \frac{v_i}{1 + \sum_{j \in S} v_j},$$

for all $t \in [T]$ and $i \in S \cup \{0\}$, where we write $v_0 := 1$. The assortment decisions of the seller are described by his/her policy: a collection of probability distributions $\pi = (\pi(\cdot | h) : h \in H)$ on \mathcal{A}_K , where

$$H := \bigcup_{t \in [T]} \{(S, Y) : Y \in S \cup \{0\}, S \in \mathcal{A}_K\}^{t-1}$$

is the set of possible histories. Then, conditionally on $h = (S_1, Y_1, \dots, S_{t-1}, Y_{t-1})$, assortment S_t has distribution $\pi(\cdot | h)$, for all $h \in H$ and all $t \in [T]$. Let \mathbb{E}_v^π be the expectation operator under policy π and preference vector $v \in \mathcal{V} := [0, \infty)^N$. The objective for the seller is to find a policy π that maximizes the total accumulated revenue or, equivalently, minimizes the accumulated regret:

$$\Delta_\pi(T, v) := T \cdot \max_{S \in \mathcal{A}_K} r(S, v) - \sum_{t=1}^T \mathbb{E}_v^\pi [r(S_t, v)],$$

where $r(S, v)$ is the expected revenue of assortment $S \subseteq [N]$ under preference vector $v \in \mathcal{V}$:

$$r(S, v) := \frac{\sum_{i \in S} v_i w_i}{1 + \sum_{i \in S} v_i}.$$

In addition, we define the worst-case regret:

$$\Delta_\pi(T) := \sup_{v \in \mathcal{V}} \Delta_\pi(T, v).$$

The main result, presented below, states that the regret of any policy can uniformly be bounded from below by a constant times \sqrt{NT} .

Theorem 1. *Suppose that $K < N/2$. Then, there exists a constant $c_1 > 0$ such that, for all $T \geq N$ and for all policies π ,*

$$\Delta_\pi(T) \geq c_1 \sqrt{NT}.$$

3. Proof of Theorem 1

3.1. Proof outline

The proof of Theorem 1 can be broken up into four steps. First, we define a baseline preference vector $v^0 \in \mathcal{V}$ and we show that under v^0 any assortment $S \in \mathcal{S}_K$ is optimal. Second, for each $S \in \mathcal{A}_K$, we define a preference vector $v^S \in \mathcal{V}$ by

$$v_i^S := \begin{cases} v_i^0(1 + \epsilon), & \text{if } i \in S, \\ v_i^0, & \text{otherwise,} \end{cases} \tag{3.1}$$

for some $\epsilon \in (0, 1]$. For each such v^S , we show that the instantaneous regret from offering a suboptimal assortment S_t is bounded from below by a constant times the number of products $|S \setminus S_t|$ not in S ; cf. Lemma 1 below. This lower bound takes into account how much the assortments S_1, \dots, S_T overlap with S when the preference vector is v^S . Third, let N_i denote the number of times product $i \in [N]$ is contained S_1, \dots, S_T , that is,

$$N_i := \sum_{t=1}^T \mathbf{1}\{i \in S_t\}.$$

Then, we use the Kullback–Leibler (KL) divergence and Pinsker’s inequality to upper bound the difference between the expected value of N_i under v^S and $v^{S \setminus \{i\}}$, see Lemma 2. Fourth, we apply a randomization argument over $\{v^S : S \in \mathcal{S}_K\}$, we combine the previous steps, and we set ϵ accordingly to conclude the proof.

The novelty of this work is concentrated in the first two steps. The third and fourth step closely follow the work of Chen and Wang [3]. These last steps are included (1) because of slight deviations in our setup, (2) for the sake of completeness, and (3) since the proof techniques are extended to the case where $K/N < 1/2$. In the work of Chen and Wang [3], the lower bound is shown for $K/N \leq 1/4$, but the authors already mention that this constraint can probably be relaxed. Our proof confirms that this is indeed the case.

3.2. Step 1: Construction of the baseline preference vector

Let $\underline{w} := \min_{i \in [N]} w_i > 0$ and define the constant

$$s := \frac{\underline{w}^2}{3 + 2\underline{w}}.$$

Note that $s < \underline{w}$. The baseline preference vector is formally defined as

$$v_i^0 := \frac{s}{K(w_i - s)}, \quad \text{for all } i \in [N].$$

Now, the expected revenue for any $S \in \mathcal{A}_K$ under v^0 can be rewritten as

$$r(S, v^0) = \frac{\sum_{i \in S} v_i^0 w_i}{1 + \sum_{i \in S} v_i^0} = \frac{s \sum_{i \in S} \frac{w_i}{w_i - s}}{K + \sum_{i \in S} \frac{s}{w_i - s}} = \frac{s \sum_{i \in S} \frac{w_i}{w_i - s}}{K - |S| + \sum_{i \in S} \frac{w_i}{w_i - s}}.$$

The expression on the right-hand side is only maximized by assortments S with maximal size $|S| = K$, in which case

$$r(S, v^0) = \max_{S' \in \mathcal{A}_K} r(S', v^0) = s.$$

It follows that *all* assortments S with size K are optimal.

3.3. Step 2: Lower bound on the instantaneous regret of v^S

For the second step, we bound the instantaneous regret under v^S .

Lemma 1. *Let $S \in \mathcal{S}_K$. Then, there exists a constant $c_2 > 0$, only depending on \underline{w} and s , such that, for all $t \in [T]$ and $S_t \in \mathcal{A}_K$,*

$$\max_{S' \in \mathcal{A}_K} r(S', v^S) - r(S_t, v^S) \geq c_2 \frac{\epsilon |S \setminus S_t|}{K}.$$

As a consequence,

$$T \cdot \max_{S' \in \mathcal{A}_K} r(S', v^S) - \sum_{t=1}^T r(S_t, v^S) \geq c_2 \epsilon \left(T - \frac{1}{K} \sum_{i \in S} N_i \right). \tag{3.2}$$

Proof. Fix $S \in \mathcal{S}_K$. First, note that since $\epsilon \leq 1$, for any $S' \in \mathcal{A}_K$, it holds that

$$\sum_{i \in S'} v_i^S \leq \frac{2s}{\underline{w} - s}. \tag{3.3}$$

Second, let $S^* \in \arg \max_{S' \in \mathcal{A}_K} r(S', v^S)$ and $\varrho^* = r(S^*, v^S)$. By rewriting the inequality $\varrho^* \geq r(S', v^S)$ for all $S' \in \mathcal{A}_K$, we find that for all $S' \in \mathcal{A}_K$

$$\varrho^* \geq \sum_{i \in S'} v_i^S (w_i - \varrho^*). \tag{3.4}$$

Let $t \in [T]$ and $S_t \in \mathcal{A}_K$. Then, it holds that

$$\begin{aligned} r(S^*, v^S) - r(S_t, v^S) &= \varrho^* - \frac{\sum_{i \in S_t} v_i^S w_i}{1 + \sum_{i \in S_t} v_i^S} \\ &= \frac{1}{1 + \sum_{i \in S_t} v_i^S} \left(\varrho^* + \sum_{i \in S_t} v_i^S \varrho^* - \sum_{i \in S_t} v_i^S w_i \right) \\ &\geq \frac{\underline{w} - s}{\underline{w} + s} \left(\varrho^* - \sum_{i \in S_t} v_i^S (w_i - \varrho^*) \right) \\ &\geq \frac{\underline{w} - s}{\underline{w} + s} \left(\sum_{i \in S} v_i^S (w_i - \varrho^*) - \sum_{i \in S_t} v_i^S (w_i - \varrho^*) \right) \end{aligned}$$

$$\begin{aligned}
 &= \frac{\underline{w} - s}{\underline{w} + s} \underbrace{\left(\sum_{i \in S} v_i^S (w_i - s) - \sum_{i \in S_t} v_i^S (w_i - s) \right)}_{(a)} \\
 &\quad - \underbrace{(\varrho^* - s) \left(\sum_{i \in S} v_i^S - \sum_{i \in S_t} v_i^S \right)}_{(b)}.
 \end{aligned}$$

Here, the first inequality is due to (3.3) and the second inequality follows from (3.4) with $S' = S$. Next, note that since $|S_t| \leq K$ and $|S| = K$, we find that

$$|S_t \setminus S| \leq |S \setminus S_t|. \tag{3.5}$$

Now, term (a) can be bounded from below as

$$\begin{aligned}
 (a) &= \sum_{i \in S \setminus S_t} v_i^S (w_i - s) - \sum_{i \in S_t \setminus S} v_i^S (w_i - s) \\
 &= \frac{s}{K} ((1 + \epsilon) |S \setminus S_t| - |S_t \setminus S|) \\
 &\geq s \frac{\epsilon |S \setminus S_t|}{K}.
 \end{aligned} \tag{3.6}$$

Here, at the final inequality, we used (3.5). Next, term (b) can be bounded from above as

$$(b) \leq \underbrace{|\varrho^* - s|}_{(c)} \underbrace{\left| \sum_{i \in S} v_i^S - \sum_{i \in S_t} v_i^S \right|}_{(d)}.$$

Now, for term (c), we note that $v_i^S \geq v_i^0$ for all $i \in [N]$. In addition, since $r(S^*, v^0) \leq s$,

$$\begin{aligned}
 \varrho^* - s &\leq \frac{\sum_{i \in S^*} v_i^S w_i}{1 + \sum_{i \in S^*} v_i^S} - \frac{\sum_{i \in S^*} v_i^0 w_i}{1 + \sum_{i \in S^*} v_i^0} \\
 &\leq \frac{1}{1 + \sum_{i \in S^*} v_i^0} \sum_{i \in S^*} (v_i^S - v_i^0) w_i \\
 &\leq \sum_{i=1}^N (v_i^S - v_i^0) = \epsilon \sum_{i \in S} v_i^0 \leq \frac{s}{\underline{w} - s} \epsilon.
 \end{aligned}$$

This entails an upper bound for (c). Term (d) is bounded from above as

$$\begin{aligned}
 (d) &\leq \sum_{i \in S \setminus S_t} v_i^S + \sum_{i \in S_t \setminus S} v_i^S \\
 &\leq (1 + \epsilon) \sum_{i \in S \setminus S_t} v_i^0 + \sum_{i \in S_t \setminus S} v_i^0 \\
 &\leq (1 + \epsilon) \frac{s}{K(\underline{w} - s)} |S \setminus S_t| + \frac{s}{K(\underline{w} - s)} |S_t \setminus S| \\
 &\leq \frac{3s}{\underline{w} - s} \frac{|S \setminus S_t|}{K}.
 \end{aligned}$$

Here, at the final inequality, we used (3.5) and the fact that $\epsilon \leq 1$. Now, we combine the upper bounds of (c) and (d) to find that

$$(b) \leq \frac{3s^2}{(\underline{w} - s)^2} \cdot \frac{\epsilon |S \setminus S_t|}{K}. \tag{3.7}$$

It follows from (3.6) and (3.7) that

$$\begin{aligned} r(S^*, v^S) - r(S_t, v^S) &\geq \frac{\underline{w} - s}{\underline{w} + s} \left(s - \frac{3s^2}{(\underline{w} - s)^2} \right) \frac{\epsilon |S \setminus S_t|}{K} \\ &\geq c_2 \frac{\epsilon |S \setminus S_t|}{K}, \end{aligned}$$

where

$$c_2 := \frac{\underline{w} - s}{\underline{w} + s} \left(s - \frac{3s^2}{(\underline{w} - s)^2} \right).$$

Note that the constant c_2 is positive if $(\underline{w} - s)^2 > 3s$. This follows from $s = \underline{w}^2 / (3 + 2\underline{w})$ since

$$(\underline{w} - s)^2 - 3s > \underline{w}^2 - s(3 + 2\underline{w}).$$

Statement (3.2) follows from the additional observation

$$\sum_{t=1}^T |S \setminus S_t| = TK - \sum_{t=1}^T |S \cap S_t| = TK - \sum_{i \in S} N_i. \quad \square$$

3.4. Step 3: KL divergence and Pinsker’s inequality

We denote the dependence of the expected value and the probability on the preference vector v^S as $\mathbb{E}_S[\cdot]$ and $\mathbb{P}_S(\cdot)$ for $S \in \mathcal{A}_K$. In addition, we write $S \setminus i$ instead of $S \setminus \{i\}$. The lemma below states an upper bound on the KL divergence of \mathbb{P}_S and $\mathbb{P}_{S \setminus i}$ and uses Pinsker’s inequality to derive an upper bound on the absolute difference between the expected value of N_i under v^S and $v^{S \setminus i}$.

Lemma 2. *Let $S \in \mathcal{S}_K$, $S' \in \mathcal{A}_K$, and $i \in S$. Then, there exists a constant c_3 , only depending on \underline{w} and s , such that*

$$KL(\mathbb{P}_S(\cdot | S') \parallel \mathbb{P}_{S \setminus i}(\cdot | S')) \leq c_3 \frac{\epsilon^2}{K}.$$

As a consequence,

$$|\mathbb{E}_S[N_i] - \mathbb{E}_{S \setminus i}[N_i]| \leq \sqrt{2c_3} \frac{\epsilon T^{3/2}}{\sqrt{K}}. \tag{3.8}$$

Proof. Let \mathbb{P} and \mathbb{Q} be arbitrary probability measures on $S' \cup \{0\}$. It can be shown, see, for example, Lemma 3 from Chen and Wang [3], that

$$KL(\mathbb{P} \parallel \mathbb{Q}) \leq \sum_{j \in S' \cup \{0\}} \frac{(p_j - q_j)^2}{q_j},$$

where p_j and q_j are the probabilities of outcome j under \mathbb{P} and \mathbb{Q} , respectively. We apply this result for p_j and q_j defined as

$$p_j := \frac{v_j^S}{1 + \sum_{\ell \in S'} v_\ell^S} \quad \text{and} \quad q_j := \frac{v_j^{S \setminus i}}{1 + \sum_{\ell \in S'} v_\ell^{S \setminus i}},$$

for $j \in S' \cup \{0\}$. First, note that by (3.3), for all $j \in S' \cup \{0\}$,

$$q_j \geq \frac{v_j^0}{1 + 2\frac{s}{\underline{w}-s}} = \frac{\underline{w}-s}{\underline{w}+s} v_j^0.$$

Now, we bound $|p_j - q_j|$ from above for $j \in S' \cup \{0\}$. Note that for $j = 0$ it holds that

$$\begin{aligned} |p_0 - q_0| &= \frac{|\sum_{\ell \in S'} v_\ell^S - \sum_{\ell \in S'} v_\ell^{S^i}|}{(1 + \sum_{\ell \in S'} v_\ell^S)(1 + \sum_{\ell \in S'} v_\ell^{S^i})} \\ &\leq |(1 + \epsilon)v_i^0 - v_i^0| = v_i^0 \epsilon. \end{aligned}$$

For $j \neq i$, since $\epsilon \leq 1$, we find that

$$|p_j - q_j| = v_j^S |p_0 - q_0| \leq 2v_j^0 v_i^0 \epsilon.$$

For $j = i$, we find that

$$\begin{aligned} |p_i - q_i| &= v_i^0 |p_0 - q_0 + \epsilon p_0| \\ &\leq v_i^0 (|p_0 - q_0| + \epsilon p_0) \\ &\leq v_i^0 (v_i^0 + 1) \epsilon. \end{aligned}$$

Therefore, we conclude that

$$\begin{aligned} \text{KL}(\mathbb{P}_S(\cdot | S') \parallel \mathbb{P}_{S^i}(\cdot | S')) &\leq \sum_{j \in S' \cup \{0\}} \frac{(p_j - q_j)^2}{q_j} \\ &\leq \frac{(p_0 - q_0)^2}{q_0} + \sum_{j \in S': j \neq i} \frac{(p_j - q_j)^2}{q_j} + \frac{(p_i - q_i)^2}{q_i} \\ &\leq \frac{\underline{w}+s}{\underline{w}-s} \left((v_i^0 \epsilon)^2 + 4(v_i^0 \epsilon)^2 \sum_{j \in S': j \neq i} v_j^0 + v_i^0 (v_i^0 + 1)^2 \epsilon^2 \right) \\ &\leq \frac{s(\underline{w}+s)}{(\underline{w}-s)^2} \left(\frac{s}{\underline{w}-s} + \frac{4s^2}{(\underline{w}-s)^2} + \left(\frac{\underline{w}}{\underline{w}-s} \right)^2 \right) \frac{\epsilon^2}{K} \\ &= c_3 \frac{\epsilon^2}{K}, \end{aligned}$$

where

$$c_3 := \frac{s(\underline{w}+s)}{(\underline{w}-s)^2} \left(\frac{s}{\underline{w}-s} + \frac{4s^2 + \underline{w}^2}{(\underline{w}-s)^2} \right).$$

Next, note that the entire probability measures \mathbb{P}_S and \mathbb{P}_{S^i} depend on T . Then, as a consequence of the chain rule of the KL divergence, we find that

$$\text{KL}(\mathbb{P}_S \parallel \mathbb{P}_{S^i}) \leq c_3 \frac{\epsilon^2 T}{K}.$$

Now, statement (3.8) follows from

$$\begin{aligned}
 |\mathbb{E}_S[N_i] - \mathbb{E}_{S \setminus i}[N_i]| &\leq \sum_{n=0}^T n |\mathbb{P}_S(N_i = n) - \mathbb{P}_{S \setminus i}(N_i = n)| \\
 &\leq T \sum_{n=0}^T |\mathbb{P}_S(N_i = n) - \mathbb{P}_{S \setminus i}(N_i = n)| \\
 &= 2T \max_{n=0, \dots, T} |\mathbb{P}_S(N_i = n) - \mathbb{P}_{S \setminus i}(N_i = n)| \\
 &\leq 2T \sup_A |\mathbb{P}_S(A) - \mathbb{P}_{S \setminus i}(A)| \\
 &\leq T \sqrt{2 \text{KL}(\mathbb{P}_S \parallel \mathbb{P}_{S \setminus i})}, \tag{3.9}
 \end{aligned}$$

where the step in (3.9) follows from, for example, Proposition 4.2 from Levin *et al.* [8] and we used Pinsker’s inequality at the final inequality. \square

3.5. Step 4: Proving the main result

With all the established ingredients, we can finalize the proof of the lower bound on the regret.

Proof of Theorem 1. Since $v^S \in \mathcal{V}$ for all $S \in \mathcal{S}_K$ and by Lemma 1, we know that

$$\begin{aligned}
 \Delta_\pi(T) &\geq \frac{1}{|\mathcal{S}_K|} \sum_{S \in \mathcal{S}_K} \Delta_\pi(T, v^S) \\
 &\geq c_2 \underbrace{\left(T - \frac{1}{|\mathcal{S}_K|} \sum_{S \in \mathcal{S}_K} \frac{1}{K} \sum_{i \in S} \mathbb{E}_S[N_i] \right)}. \tag{3.10}
 \end{aligned}$$

We decompose (a) into two terms:

$$(a) = \underbrace{\frac{1}{|\mathcal{S}_K|} \sum_{S \in \mathcal{S}_K} \frac{1}{K} \sum_{i \in S} \mathbb{E}_{S \setminus i}[N_i]}_{(b)} + \underbrace{\frac{1}{|\mathcal{S}_K|} \sum_{S \in \mathcal{S}_K} \frac{1}{K} \sum_{i \in S} (\mathbb{E}_S[N_i] - \mathbb{E}_{S \setminus i}[N_i])}_{(c)}.$$

Recall that $c = K/N \in (0, 1/2)$. By summing over $S' = S \setminus i$ instead of over S , we bound (b) from above by

$$(b) = \frac{1}{|\mathcal{S}_K|} \sum_{S' \in \mathcal{S}_{K-1}} \frac{1}{K} \sum_{i \notin S'} \mathbb{E}_{S'}[N_i] \leq \frac{|\mathcal{S}_{K-1}|}{|\mathcal{S}_K|} T \leq \frac{c}{1-c} T,$$

where the first inequality follows from $\sum_{i \in [N]} \mathbb{E}_{S'}[N_i] \leq TK$, and the second inequality from

$$\frac{|\mathcal{S}_{K-1}|}{|\mathcal{S}_K|} = \frac{\binom{N}{K-1}}{\binom{N}{K}} = \frac{K}{N-K+1} \leq \frac{K/N}{1-K/N}.$$

Now, (c) can be bounded by applying Lemma 2:

$$(c) \leq \sqrt{2c_3} \frac{\epsilon T^{3/2}}{\sqrt{K}} = \frac{\sqrt{2c_3} \epsilon T^{3/2}}{\sqrt{c} \sqrt{N}}.$$

By plugging the upper bounds on (b) and (c) in (3.10), we obtain

$$\begin{aligned} \Delta_\pi(T) &\geq c_2\epsilon \left(T - \frac{c}{1-c}T - \frac{\sqrt{2c_3}}{\sqrt{c}} \frac{\epsilon T^{3/2}}{\sqrt{N}} \right) \\ &= c_2\epsilon \left(\frac{1-2c}{1-c}T - \frac{\sqrt{2c_3}}{\sqrt{c}} \frac{\epsilon T^{3/2}}{\sqrt{N}} \right). \end{aligned}$$

Now, we set ϵ as

$$\epsilon = \min \left\{ 1, \frac{(1-2c)\sqrt{c}}{2(1-c)\sqrt{2c_3}} \sqrt{N/T} \right\}.$$

This yields, for all $T \geq N$,

$$\Delta_\pi(T) \geq \min \left\{ \frac{c_2\sqrt{2c_3}}{\sqrt{c}} T, \frac{c_2(1-2c)^2c}{8(1-c)\sqrt{2c_3}} \sqrt{NT} \right\}.$$

Finally, note that for $T \geq N$ it follows that $T \geq \sqrt{NT}$ and therefore

$$\Delta_\pi(T) \geq c_1\sqrt{NT},$$

where

$$c_1 := \min \left\{ \frac{c_2\sqrt{2c_3}}{\sqrt{c}}, \frac{c_2(1-2c)^2c}{8(1-c)\sqrt{2c_3}} \right\} > 0. \quad \square$$

Acknowledgments. The authors thank the Editor in Chief and an anonymous reviewer for their positive remarks and useful comments.

Competing interest. The authors declare no conflict of interest.

References

[1] Agrawal, S., Avadhanula, V., Goyal, V., & Zeevi, A. (2017). Thompson sampling for the MNL-bandit. In *Conference on Learning Theory (COLT)*, pp. 76–78.

[2] Agrawal, S., Avadhanula, V., Goyal, V., & Zeevi, A. (2019). MNL-bandit: a dynamic learning approach to assortment selection. *Operations Research* 67(5): 1453–1485.

[3] Chen, X. & Wang, Y. (2018). A note on a tight lower bound for MNL-bandit assortment selection models. *Operations Research Letters* 46(5): 534–537.

[4] Chen, X., Wang, Y., & Zhou, Y. (2018). An optimal policy for dynamic assortment planning under uncapacitated multinomial logit models. *ArXiv e-print*. <https://arxiv.org/abs/1805.04785>.

[5] Cheung, W.C. & Simchi-Levi, D. (2017). Assortment optimization under unknown multinomial logit choice models. *ArXiv e-print*. <https://arxiv.org/abs/1704.00108>.

[6] Farias, V.F., Jagabathula, S., & Shah, D. (2013). A nonparametric approach to modeling choice with limited data. *Management Science* 59(2): 305–322.

[7] Kallus, N. & Udell, M. (2016). Dynamic assortment personalization in high dimensions. *ArXiv e-print*. <https://arxiv.org/abs/1610.05604>.

[8] Levin, D.A., Peres, Y., & Wilmer, E.L. (2017). *Markov chains and mixing times*, 2nd ed. Providence, RI: American Mathematical Society.

[9] Rusmevichientong, P., Shen, Z.-J.M., & Shmoys, D.B. (2010). Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research* 58(6): 1666–1680.

[10] Sauré, D. & Zeevi, A. (2013). Optimal dynamic assortment planning with demand learning. *Manufacturing & Service Operations Management* 15(3): 387–404.

Cite this article: Peeters Y and den Boer AV (2022). A regret lower bound for assortment optimization under the capacitated MNL model with arbitrary revenue parameters. *Probability in the Engineering and Informational Sciences* 36, 1266–1274. <https://doi.org/10.1017/S0269964821000395>