A Statistical Learning Approach to Personalization in Revenue Management

Zachary Owen, David Simchi-Levi, Clark Pixton, Xi Chen

A Statistical Learning Approach to Personalization in Revenue Management

Abstract

We develop a general framework for modeling decision problems in which actions can be personalized by taking into account the information available to the decision maker at each stage through binomial and multinomial logistic regression. We demonstrate the application of our method to customized pricing decisions and personalized assortment optimization. We show that learning under our model takes place reliably by establishing finite-sample high probability convergence guarantees for model parameters which hold regardless of the number of customer types, which can be potentially uncountable. The parameter convergence guarantees can then be extended to performance bounds in operational problems in which the gap between the expected revenue of our suggested method and the optimal personalized decisions shrinks in proportion to the square root of the size of the training set. Tests of our method on real transaction data for airline seating reservations demonstrate that effective customized pricing can increase revenue by at least 7% over the best single-price. Further tests on simulated data suggest that the methods we developed tend to perform better than the worst-case bounds we present.

Keywords: machine learning; statistics; price discrimination; assortment optimization; logistic regression; choice modeling; revenue management

1. Introduction

The increasing prominence of electronic commerce has given businesses an unprecedented ability to understand their customers as individuals and to tailor their services for them appropriately. This benefit is two-fold: customer profiles and data repositories often provide information that can be used to predict which products and services are most relevant to a customer, and the fluid nature of electronic services allows for this information to be used to optimize their experience in real-time, see Murthi and Sarkar (2003). For instance, Linden et al. (2003) document how Amazon.com has used personalization techniques to optimize the selection of products it recommends to users for many years, dramatically increasing click-through and conversion rates as compared to static sites. Other companies, such as Netflix, have implemented personalization through recommender systems as described by Amatriain (2013) to drive revenue indirectly by improving customer experience.

To implement a personalization strategy it has been widely proposed to divide the customer base into distinct segments and to tailor the service to each type appropriately as in Chan et al. (2011). This segmentation of customers is often accomplished in practice by dividing them into archetypical categories based on broad, easily observable characteristics such as business versus leisure travelers in the case of the airline industry as discussed in Talluri and van Ryzin (2004b). Such methods can increase revenue over models in which customers are assumed to be homogeneous, but are likely to be suboptimal in the presence of large amounts of data.

In this paper, we consider a statistical model and algorithm that use observed contextual information, rather than previously defined customer segments, to inform decisions. This approach represents a shift from thinking about personalization in terms of customer types towards personalized management decisions as a function of the unique relevant information available at the time of each decision. In the language of machine learning, we cast personalized decision making as a supervised learning problem, where past transactional data is used to discover the underlying relationship between contextual information and customer behavior, and predictions are made based on this relationship. This approach takes full advantage of available data in at least two ways. First, it allows the seller to consider more complex relationships between context and customer behavior than can be captured with just a few customer segments. Our approach also allows previous learning to be generalized easily, even to previously unseen types, allowing for customization even in the case of novel customer data.

As a demonstration of the advantages of this approach, consider again the example of business and leisure travelers in the airline industry. Within the business customer segment, there are large businesses with many assets and low price sensitivity, but also small firms that may not yet be financially established. Similarly, leisure travelers are often more price sensitive, but there are some wealthy travelers who may behave more like business customers. Transaction-specific contextual information is often rich enough to capture these "second-order" trends within the broad categories, which can then be leveraged to drive incremental revenue or to improve customer experience. With the granularity of information available today, it is likely that each new customer represents a unique pattern of information, and our personalization strategy should be flexible enough to both learn from and optimize for the full diversity of the customer population.

To accomplish this type of strategy, we model demand or customer choice with a binomial or multinomial logit function, where the arguments to the logit are assumed to be linear functions of observed features. Since logistic regression models have been well-studied and are very popular in practice, this approach leads both to theoretical guarantees and to a practical, data-driven algorithm. Our main contributions are summarized here.

Incorporation of actions of the seller into estimable models of choice. In many applications of statistics and machine learning, the data of interest consist of contextual information (or features) and outcomes. In our case, however, there is a third element to the data - the action of the seller. This action may be a price set by a firm or an assortment offered to a customer, and it has an important effect on the observed outcome of the transaction. We provide ways to incorporate these actions into learning algorithms for customer behavior models.

Estimation of personalized choice behavior with transaction data. By assuming a parametric form of the choice model, we allow for estimation of customer choice across arbitrary types, even for customers with feature vectors not before observed. It is natural to expect this when customer features consist of continuous variables and several or many dimensions. An example is an airline ticket sale, in which the features for the transaction may include time remaining until the flight takes off, origin, destination, an indicator for weekend versus weekday purchases, customer web browser, etc. Our framework allows for all of this information to be taken into account in the model of each customer's choice, and provides a practical way of estimating all necessary parameters using transaction data of the form available to modern firms.

Statistical revenue bounds. Using a probabilistic analysis of the maximum likelihood estimates for logistic regression, we provide a parameter bound in terms of the number of samples. This result allows us to bound the expected revenue gap between our algorithm's performance and the performance of an oracle, and demonstrates that this revenue gap shrinks to zero as the number of samples grows large. A version of these bounds also holds in the high-dimensional case where the dimension of the feature space grows with the number of samples.

Good performance with real transaction data. We test the performance of our proposed algorithm specialized for customized pricing on transaction data from a European airline and show a significant increase in revenue over the best single price policy. We also show that our algorithms perform well with simulated data for customized pricing.

Throughout this paper we demonstrate our analysis by examining the application of our approach to two problems in operations management: customized pricing and personalized assortment optimization. In customized pricing a seller aims to offer her product to each customer at a price that maximizes expected profit, given information she has obtained about the transaction and the customer. Customized pricing, or price discrimination, is quite common in business to business transactions. Despite its limited application to general settings due to customer satisfaction and legal issues, it is predicted that the practice will spread and become more widely accepted as more data becomes available (see Golrezaei et al. (2014)). Our work in this application builds on the ideas and work of Carvalho and Puterman (2005).

In personalized assortment optimization a seller aims to show to each customer the assortment of products which maximizes her expected profit for each individual customer arrival, again given the contextual information. In the traditional brick-and-mortar context, implementation of a personalized assortment strategy is out of the question due to prohibitive setup costs. However, for online retailers personalized assortments are possible and even natural.

We will demonstrate that our approach makes a significant contribution in both of these growing areas of application. We also note that our style of analysis is not limited to these specific instances, and can be applied in many other contexts such as online advertisement allocation, crowdsourcing task assignment, personalized medicine, and other problems in which personalization can aid in optimal decision-making.

2. Literature Review

Our work combines two recent themes within operations management: learning problem parameters from data and personalizing decisions using contextual information. One recent paper that incorporates both of these aspects is Rudin and Vahn (2014), in which the authors consider a "big data newsvendor" problem. In their model, the decision maker has access to relevant information about the current environment before making each order quantity decision, and aims to find not just one optimal order quantity, but an optimal *rule* for mapping contextual information into an order decision. They propose techniques for choosing such a rule using only observed data and prove bounds on the cost that depend on the amount of past data available. Tools from machine learning provide the main machinery for their proofs. They also test their approach on real data for a nurse staffing problem and demonstrate a remarkable improvement over approaches which do not take contextual information into account. Our work follows in this theme by examining how pricing, assortment, and other customer-facing decisions can be personalized using contextual information and past data, resulting in revenue bounds that improve as the amount of data available increases.

The main applications of personalization that we consider here are customized pricing and personalized assortment optimization. As both pricing and assortment optimization have been well studied recently, we provide a brief overview of both of these areas. We highlight in each application area recent work focusing on learning and personalization.

Recently, dynamic pricing and demand learning has been a popular theme in the field of revenue management and a brief survey of some early work in this field can be found in Aviv et al. (2012). In these settings it is often assumed that the demand function of the entire population is unknown, but it is possible to obtain information about the structure of demand through price experimentation. A common modeling assumption, (as used in Broder and Rusmevichientong (2012), for example), is that the true market demand function is specified by a parametric choice model. Using price experimentation, they develop a pricing policy that achieves the minimum possible asymptotic regret in comparison to a clairvoyant who knows the full demand model. Even more related to the current work is that of Carvalho and Puterman (2005), who assume a logistic regression model for demand as a function of price. They mention maximum likelihood estimation for their model, which we study here in detail.

At least in the field of operations research and management science (as opposed to economics), literature on personalized pricing is sparse. This is likely due to the fact that pure price discrimination is often thought to have limited application. However, there have been some important contributions. In Carvalho and Puterman (2005), the authors mention that customer-specific features may be included as part of their model. Aydin and Ziya (2009) consider the case of customized pricing in which customers belong to a high or low reservation price group and provide a signal to the seller that provides some information as to how likely they are to belong to the higher price group. Among other results, they develop conditions on the relationship between the signal and a customer's probability of belonging to the higher price group under which the optimal price to offer is monotonic in the signal strength. Netessine et al. (2006) consider a form of personalized dynamic pricing in their treatment of cross-selling based on the other items that each consumer is considering purchasing.

In many other cases, models of price discrimination are actually cast as multi-product models, where the different price levels come with different qualifications and extras as in the airline industry. See Talluri and van Ryzin (2004a) and Belobaba (1989) for examples of this type. While allowing for customized pricing for a single product based on customer attributes, which would be appropriate in the insurance industry and in business to business transactions, our conception of features also allows for dynamic pricing of differing products based on their individual attributes, the effects of which may be learned over time. For example, in our work with a corporate partner, the price of seating reservation privileges can be customized to features of the generic transaction such as time of day in which the website was accessed and details concerning the flight itself such as the origin and destination pair and the time of departure.

Assortment optimization in the static case was brought to the attention of the operations management community by van Ryzin and Mahajan (1999). Since that time assortment optimization techniques and models have been heavily researched, with much past work well summarized in Kök et al. (2008). Some setups such as those of Talluri and van Ryzin

(2004a) and Golrezaei et al. (2014) allow for a general model of customer choice, but others study structural properties specific to certain choice models. The multinomial logit model (MNL), our model in the current work, is among the most commonly studied models of customer choice for assortment optimization. Some examples include Rusmevichientong et al. (2010) and Rusmevichientong et al. (2014).

One notable paper that specifically treats learning of parameters for choice models in assortment optimization and related problems is Vulcano et al. (2008), which presents an algorithm for estimating true demand from censored transaction data and proves convergence of the algorithm. Also, in Rusmevichientong et al. (2010), the authors consider a constrained version of assortment optimization and propose a dynamic algorithm that incorporates learning the parameters of the MNL model.

As discussed above, personalized assortment optimization has only been considered recently assortment planning literature. Bernstein et al. (2011) approached this problem by assuming multiple customer types, each with its own choice model. The authors show properties of an optimal assortment policy for two products in the presence of inventory considerations. Golrezaei et al. (2014) also consider multiple customer types, providing a practical algorithm for personalization under inventory constraints and proving a strong worst-case performance bound. They also discuss the issue of estimating model parameters, and show that a version of their worst-case competitive ratio bound still holds for the case when parameters are estimated from the data.

The rest of the paper will proceed as follows. We give our approach and models in Section 3. In Section 4 we present the algorithms for customized pricing and assortment optimization. Sections 5 and 6 are devoted to proving revenue bounds for the two problems under various assumptions, including a high-dimensional result. In Section 7 we show the results of experiments on both real and simulated data.

3. A General Model

In this section we present a general modeling framework for data-driven decision problems that include decision-specific context information, or features. We consider two applications in detail, customized pricing and assortment optimization, but we note that our approach is certainly not limited to these domains.

A decision maker observes a vector of features $z \in \mathcal{Z} \subseteq \mathcal{R}^d$ that encodes information about the context of the specific decision at hand. Taking into account z, he chooses an action a from a problem-specific action space $\mathcal{A}$. After the decision has been made, he observes an outcome y from a finite set $\mathcal{Y}$ and gains a random reward from a finite set $\mathcal{R}$. The probability of outcome y depends on the context z and the decision a. The reward $r_a(y)$ for outcome y may also depend on the decision a. He would like to make the decision that maximizes his expected reward given the context z. The key to our modeling framework is a method for capturing the interaction between features, decision, and outcome using a binomial or multinomial logit model, depending on the application. These models give conditional outcome probabilities $\mathbb{P}_z(y;a)$. We will derive the specific form for outcome probabilities as a function of features and decision for both customized pricing and assortment optimization later in the section.

Given the outcome probabilities from the problem-specific logit model, we can write the expected reward:

$$f_z(a) = \sum_{y\in\mathcal{Y}} r_a(y)\mathbb{P}_z(y;a).$$

The algorithm we present estimates the expected reward and then maximizes over all possible decisions. Before stating the algorithms in detail, we give specific models for customized pricing and personalized assortment optimization.

3.1. Customized Pricing Model

The first application of our model is in the case where a seller has a single product without inventory constraints and wishes to offer a price that will maximize his revenue. In basic form, the single-product pricing problem consists of a set $\mathcal{A} = \{p^1,\dots,p^K\}$ of candidate prices and a probability of purchase $\mathbb{P}(y=1;p^k),\ k=1,\dots,K$. Here the outcome y is a binary decision and is equal to one if the customer purchases the product and zero otherwise. Thus, the expected revenue function is

$$f(p) = p\mathbb{P}(y=1;p). \tag{1}$$

Without any other information, the seller would maximize $f(\cdot)$ over $\mathcal{A}$.

In the customized pricing problem, we assume that the seller has the ability to offer a different price $p\in\mathcal{A}$ to each customer and that each customer is associated with a feature vector $z\in\mathcal{Z}\subseteq\mathbb{R}^m$ which supplies the context for each pricing decision. This vector z is observed by the seller before choosing a price.

This information structure allows us to define a personalized demand function $\mathbb{P}_z(y=1;p)$, which is the probability that a customer with the feature vector z purchases the product at price p. We model outcome probabilities $\mathbb{P}_z(y=1;p)$ using a logistic regression model, introducing parameters β^* and γ^*:

$$\log \frac{\mathbb{P}_z(y=1;p,\beta^*,\gamma^*)}{1-\mathbb{P}_z(y=1;p,\beta^*,\gamma^*)} = \sum_{k=1}^{K} \beta_k^* \mathbb{I}(p=p^k) + \sum_{j=1}^{m} \gamma_j^* z_j$$

This model gives a specific form for the demand function:

$$\mathbb{P}_z(y=1;p,\beta^*,\gamma^*) = \frac{1}{1+\exp\left(-\left(\sum_{k=1}^{K} \beta_k^* \mathbb{I}(p=p^k) + \sum_{j=1}^{m} \gamma_j^* z_j\right)\right)},$$

where $\mathbb{I}(A)$ is the indicator function which takes the value 1 when the event A is true and 0 otherwise. For simplicity of notation, let $x = (\mathbb{I}(p=p^1),\ldots,\mathbb{I}(p=p^K),z) \in \mathbb{R}^{K+m}$ and let $\theta^* = (\beta^*,\gamma^*) \in \mathbb{R}^{K+m}$ be the true parameter vector. We can then write the model $\mathbb{P}_z(y=1;p,\beta^*,\gamma^*)$ in a more succinct way:

$$\mathbb{P}_z(y=1;p,\beta^*,\gamma^*) = \sigma(\langle x,\theta^*\rangle),$$

where $\sigma(c) = \frac{1}{1+e^{-c}}$ is the sigmoid function.

The reward gained for each customer given the decision to offer price p is p with probability $\mathbb{P}_z(y=1;p,\beta^*,\gamma^*)$, and zero otherwise. Thus, the expected reward is given by

$$f_z(p,\theta^*) := p\ \mathbb{P}_z(y=1;p,\beta^*,\gamma^*). \tag{2}$$

We pause to make explicit the connection between customized pricing notation and the general notation. For customized pricing, each action a is a price $p \in \mathcal{A} = \{p^1,\ldots,p^K\}$. As mentioned above, outcomes $y \in \{0,1\}$ correspond to purchase decisions, and reward $r_a(y) = r_p(y) = p$ if $y=1$ and zero if $y=0$. Using these mappings one can identify $\mathbb{P}_z(y=1;p,\beta^*,\gamma^*)$ and $f_z(p,\theta^*)$ as the problem-specific versions of $\mathbb{P}_z(y;a)$ and $f_z(a)$.

We also note that our limitation of the price effect to the vector β^* is simply for clarity of exposition and to highlight the use of price as a feature in our model. In practice, this model also applies to estimating interaction effects between offered prices and other features. We have found such interaction effects to be especially useful in our work with real transaction data.

3.2. Personalized Assortment Optimization Model

Our framework also applies to the challenging problem of personalized assortment optimization. For background, in the original assortment optimization problem the decision maker has J products indexed by $\{1,2,\ldots,J\}$. For each product j, let r_j be its associated revenue, fixed a priori. We also have another "no-purchase" option indexed by zero with $r_0 = 0$. Without loss of generality, we assume that the products are indexed such that $r_J \geq r_{J-1} \geq \ldots \geq r_1 \geq r_0 = 0$. The decision maker must choose an assortment $S \in \mathcal{A}$ to show to the customer, where $\mathcal{A}$ is some set of feasible assortments.

In the general form of the problem, we assume that customers choose among the products according to some probabilities $\mathbb{P}(j;S)$, that is, the probability that a customer chooses product j given that she was shown assortment S is equal to $\mathbb{P}(j;S)$. The expected revenue for a given assortment (without any contextual information) can then be written as

$$f(S) = \sum_{j \in S} r_j \mathbb{P}(j;S), \tag{3}$$

and the decision maker then maximizes $f(S)$ over $S \in \mathcal{A}$.

The problem so far puts very few constraints on the choice model of the customer. One way to practically estimate the probabilities $\mathbb{P}(j;S)$ is to make further assumptions. In particular, we assume the customers choose among the offered products according to random utility maximization, where customers have a utility

$$U_j = V_j + \epsilon_j$$

for each product j. Here, ϵ_j is a standard Gumbel random variable with mean zero and we view V_j as the mean utility of product j. We normalize the utility of the no purchase option to zero.

If we offer the assortment $S \in \mathcal{A}$ of products to the customers, then a customer chooses the product with the highest utility if the utility of this product is positive, but otherwise, leaves without purchasing anything. It is a standard result in discrete choice theory that if we offer the assortment S to the customers, then a customer chooses product $j \in S$ with probability

$$\mathbb{P}(j;S) = \frac{e^{V_j}}{1 + \sum_{l \in S} e^{V_l}}. \tag{4}$$

The choice model above is known as the multinomial logit (MNL) model, and is widely used to model discrete choice. It is an extension of the logit model to cases with more than two alternatives. Inserting the MNL choice probabilities into (3) gives an expected revenue function in terms of the parameters V_j, $j \in [J]$, which can be estimated from data.

For the personalized problem, suppose that before choosing an assortment S, the decision maker observes a vector of features $z \in \mathcal{Z} \subseteq \mathbb{R}^d$. We can modify our assumptions on the choice probabilities to take into account this new information. In particular, we assume that each feature vector z corresponds to a different utility for each product j:

$$U_j^z = V_j^z + \epsilon_j,$$

where ϵ_j is again a standard Gumbel random variable and we view V_j^z as the mean utility of product j *for the customer with feature vector* z. We assume the following form on the utility:

$$V_j^z = \langle \theta_j^*, z \rangle,$$

where $\theta_j^* \in \mathbb{R}^d$ for $1 \leq j \leq J$. This reduces the parameter space from an arbitrary set to a dJ-dimensional linear space, and thus allows us to generalize learning from previous customers to future customers with new feature vectors. With this structural assumption, we can modify (4) for the personalized case:

$$\mathbb{P}_z(j; S, \theta^*) = \frac{e^{V_j^z}}{1 + \sum_{l \in S} e^{V_l^z}} = \frac{\exp\{\langle \theta_j^*, z \rangle\}}{1 + \sum_{l \in S} \exp\{\langle \theta_l^*, z \rangle\}}. \tag{5}$$

If $\theta^* = (\theta_1^*, \ldots, \theta_m^*) \in \mathbb{R}^{dJ}$ fully specifies customer behavior, and we offer the assortment S_z to the customer with the feature vector z, then the expected revenue obtained from this customer can be written as,

$$f_z(S, \theta^*) = \sum_{j \in S} r_j \mathbb{P}_z(j; S, \theta^*). \tag{6}$$

As in the previous subsection, we note the mapping between the general framework notation and the notation specific to the personalized assortment optimization problem. The action a in this case is the subset S, with action space $\mathcal{A}$ being the set of feasible assortments. The possible outcomes y are products, with $y \in \{0, \ldots, J\}$, and $r_a(y) = r_S(j)$ and is equal to r_j if $j \in S$ and zero otherwise. Using these mappings one can identify $\mathbb{P}_z(j; S, \theta^*)$ and $f_z(S, \theta^*)$ as the problem-specific versions of $\mathbb{P}_z(y; a)$ and $f_z(a)$.

Now that we have some concrete examples of our approach, we will highlight some of its important characteristics. By creatively defining features, our framework captures the interaction between decisions and outcomes within the logit model. Because we assume that past feature-decision-outcome data are observed and recorded, these models give rise to practical algorithms via maximum likelihood estimation for logistic regression, which is well-understood and widely used for statistics and machine learning problems.

In addition to motivating good algorithms, our approach leads to important theoretical bounds on revenue. We extend theory about maximum likelihood estimates to bound the revenue gap between the algorithm's decision and the best decision with high probability in terms of the number of samples available. This theory can also be extended to the high-dimensional setting where the number of features grows with the number of samples.

4. Algorithm

We assume the decision maker has access to a set $\mathcal{T} = \{(z_1, a_1, y_1), \ldots (z_n, a_n, y_n)\}$ of n samples of past features, decisions, and outcomes. We can calculate the negative log-likelihood

$$\ell_n(\mathcal{T}, \theta) = -\frac{1}{n} \sum_{i=1}^{n} \log\left(\mathbb{P}_{z_i}(y_i; a_i, \theta)\right).$$

Also, recall from (3) that $f_z(a, \theta)$ is the expected revenue of decision a given features z and parameter vector θ. The Personalized Revenue Maximization Algorithm is given below, assuming that we can pre-determine a positive number R such that $\|(\theta^*)\|_1 \leq R$. In practice, one can either tune this R for better performance or fix a large enough number R.

For the customized pricing problem, the actions a_i correspond to offered prices p_i, and the outcomes y_i are binary indicators of purchase decisions. The problem-specific negative log-likelihood can be calculated as follows:

$$\begin{aligned}
\ell_n(\mathcal{T}; \beta, \gamma) &= -\frac{1}{n} \sum_{i=1}^{n} \left[y_i \log \mathbb{P}_{z_i}(p_i; \beta, \gamma) + (1 - y_i) \log(1 - \mathbb{P}_{z_i}(p_i; \beta, \gamma))\right] \\
&= -\frac{1}{n} \sum_{i=1}^{n} \left[y_i \log \frac{\mathbb{P}_{z_i}(p_i; \beta, \gamma)}{1 - \mathbb{P}_{z_i}(p_i; \beta, \gamma)} + \log\left(1 - \mathbb{P}_{z_i}(p_i; \beta, \gamma)\right)\right]
\end{aligned}$$

Personalized Revenue Maximization Algorithm

1. Fit the regularized logistic regression on the observed data:

$$(\widehat{\theta}) = \underset{\|(\theta)\|_1 \le R}{\arg\min}\, \ell_n(\mathcal{T}, \theta)$$

2. Obtain the estimate of outcome probability $\mathbb{P}_z(y; a, \widehat{\theta})$ for every $z \in \mathcal{Z}$ and $a \in \mathcal{A}$.
3. Construct the decision policy $h : \mathcal{Z} \to \mathcal{A}$ as $h(z) = a^*$ where

$$a^* = \underset{a \in \mathcal{A}}{\arg\max}\, f_z(a, \widehat{\theta}). \tag{7}$$

$$\begin{aligned} &= \frac{1}{n}\sum_{i=1}^{n}\Bigg[-y_i\Big(\sum_{k=1}^{K}\beta_k \mathbb{I}(p_i = p^k) + \sum_{j=1}^{m}\gamma_j z_{ij}\Big) \\ &\qquad + \log\Big(1 + \exp\Big(\sum_{k=1}^{K}\beta_k \mathbb{I}(p_i = p^k) + \sum_{j=1}^{m}\gamma_j z_{ij}\Big)\Big)\Bigg] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[-y_i\langle x_i, \theta\rangle + \log\left(1 + \exp\left(\langle x_i, \theta\rangle\right)\right)\right], \end{aligned}$$

and the expected reward function is given in (2). In summary, the algorithm begins by using regularized logistic regression to learn a relationship between the context as well as the offered price and the probability of purchase. It then proceeds to select the price that maximizes expected revenue under this model.

In personalized assortment optimization, each action a_i is a subset of products S_i shown to the customer in transaction i. The outcomes y_i are the purchase decision: either a product index or a no-buy decision, indexed by zero. The negative log-likelihood is

$$\ell_n(\mathcal{T}, \theta) = \frac{1}{n}\sum_{i=1}^{n}\left[-\langle \theta_{j_i}, z_i\rangle + \log\left(1 + \sum_{l \in S_i}\exp\{\langle \theta_l, z_i\rangle\}\right)\right], \tag{8}$$

and the expected reward can be found in (6). In this case, the algorithm begins by regularized logistic regression to learn the relationship between the context as well as the selected assortment and the probability of purchase for each item. If there are J products available to offer, it then proceeds to select the assortment $\hat{S}$ from some set of feasible subsets $\mathcal{A}$ that maximizes the expected revenue given the model.

We note that when the point estimates of each parameter vector $\widehat{\theta}_j$ are used in this final step and $\mathcal{A}$ is taken to be all subsets of J products, the maximization over the power set of subsets of products can be reduced to maximization over *revenue-ordered assortments*, that is subsets of the form $S_k = \{1, \ldots, k\}$ for some $k \in \{1, 2, \ldots, J\}$, as demonstrated in Talluri and van Ryzin (2004a). This is because under this estimate the mean utility V_j^z is taken to be deterministic. Further, since we assume that customer behavior is specified by the true parameter vectors θ_j^*, it follows that the optimal such set is also a revenue-ordered assortment. Thus this property ensures that our maximization scales linearly with number of products J and that both our estimated optimal assortment and the true optimal assortment belong to the same limited subset of $\mathcal{A}$. This reduction of the search space is not trivial. By contrast, if the parameters V_j^z were random then as demonstrated by Rusmevichientong et al. (2014), the maximization in step three above is NP-complete even when the distributions of the model parameters are known exactly. Other papers, such as Rusmevichientong et al. (2010), consider the maximization with constraints on $\mathcal{A}$ and give tractable methods of optimization.

5. Theory: Well-specified Model Setting

With some assumptions on the inputs and outputs of the problem at hand, we will prove bounds on the optimality gap of the algorithm above in terms of the number of samples n. Our analysis focuses on the case of binomial logistic regression as used in customized pricing, but also holds for multinomial logistic regression problems.

For the first part of the analysis, we assume that there exists θ^* such that for all observed data $i \in [n]$,

$$\mathbb{P}(y_i) = \mathbb{P}_{z_i}(y_i; a_i, \theta^*), \tag{9}$$

which means the logistic model is the correct underlying model of outcome probabilities. We use as our benchmark the oracle policy that knows this true θ^*. Let $\widehat{a}$ be the action recommended by the algorithm in (7). Using properties of the maximum likelihood estimates, for any feature vector z we aim to bound the optimality gap

$$f_z(a^*, \theta^*) - f_z(\widehat{a}, \theta^*).$$

5.1. Customized Pricing Bound

As mentioned above, we will give the detailed revenue bound proof for the case of the customized pricing problem. In the context of this application, (9) can be specialized to say that

$$\mathbb{P}(y_i = 1) = \mathbb{P}_{z_i}(y_i = 1; p_i, \beta^*, \gamma^*) = \frac{1}{1 + \exp\left(-\left(\sum_{k=1}^{K} \beta_k^* \mathbb{I}(p_i = p^k) + \sum_{j=1}^{m} \gamma_j^* z_{ij}\right)\right)}, \tag{10}$$

where z_{ij} is the j-th component of the feature vector corresponding to transaction i.

Recall that $x_i = (\mathbb{I}(p_i = p^1), \ldots, \mathbb{I}(p_i = p^K), z_i)$. In addition to the well-specified model assumption, we assume that the outputs $y_i \in \{0,1\}$ are independent given each x_i. We also assume bounded inputs, i.e., there exists a constant $B' > 0$, such that for any $i \in [n]$ and $j \in [m]$

$$|z_{ij}| \leq B',$$

which further implies that for any $i \in [n]$ and $j \in [d]$,

$$|x_{ij}| \leq \max(B', 1) \triangleq B. \tag{11}$$

This assumption guarantees that the data we train on will not contain arbitrarily large elements which could have an outsize effect on our learning procedure.

We consider both deterministic design, where the input feature vectors z_i are viewed as fixed quantities and only outputs are random, and random design, where inputs z_i and outputs y_i are both randomly drawn from some distribution. The remainder of our assumptions differ between these two cases.

1. In the deterministic design setting, we assume that

$$\lambda_{\min}(\Sigma_n) \geq \frac{\rho}{2} > 0, \tag{12}$$

   where $\Sigma_n = \frac{1}{n}\sum_{i=1}^n x_i x_i^T$.

2. For the random design, we assume that $\{x_i\}_{i=1}^n$ are i.i.d. and each feature vector z_i is a sub-Gaussian random vector with sub-Gaussian norm ψ_z, i.e.,

$$\psi_z = \sup_{\|w\|_2 \leq 1} \|\langle z, w\rangle\|_{\psi_2},$$

   where z has the same distribution as any z_i. Under this assumption, it is easy to show that each x_i is also a sub-Gaussian random vector with sub-Gaussian norm

$$\psi_x \leq \psi_z + 1 \triangleq \psi.$$

   To see this, let $v_i = \big(\mathbb{I}(p_i = p^1), \ldots, \mathbb{I}(p_i = p^K)\big)$ and v be the vector which has the same distribution as each v_i. Then

$$\begin{aligned}
\psi_x &= \sup_{\|(w_v, w_z)\|_2 \leq 1} \|\langle v, w_v\rangle + \langle z, w_z\rangle\|_{\psi_2} \\
&\leq \sup_{\|w_v\| \leq 1} \|\langle v, w_v\rangle\|_{\psi_2} + \sup_{\|w_z\|_2 \leq 1} \|\langle z, w_z\rangle\|_{\psi_2} \\
&\leq \psi_z + 1.
\end{aligned}$$

Further, we assume that

$$\lambda_{\min}(\Sigma) > \rho > 0, \tag{13}$$

where $\Sigma = \mathbb{E}(xx^T)$.

We note that the strictly positive smallest eigenvalue assumptions ensure that each feature provides sufficently unique information. In the random design setting we also make the assumption of a sub-Gaussian distribution of the feature vectors which ensures that the tails of the distribution are sufficiently well behaved to enable effective learning. Many common distributions are sub-Gaussian, including any bounded distribution, and so the assumption is not too restrictive and is used to simplify the analysis. It is widely used in statistics research (see Bühlmann and van de Geer (2011), for example).

Let $d = K + m$. In the low dimensional setting where $d \ll n$, we prove that under the assumptions above, with probability, at least $1 - \frac{2}{n}$,

$$\|\widehat{\theta} - \theta^*\|_2 \leq \frac{C_{cp}(R, B, \psi)}{\rho} \sqrt{\frac{d \log(nd)}{n}}, \tag{14}$$

where $C_{cp}(R, B, \psi)$ is a constant only depending on R, B and ψ. This will drive our revenue bound. To prove (14), we first establish the strong convexity of the loss ℓ_n with strong convexity parameter $\eta > 0$. In our proofs, we suppress the data argument $\mathcal{T}$ in the function ℓ_n for convenience. Let $\widehat{\Delta} = \widehat{\theta} - \theta^*$ denote the error in our estimate of the true parameter vector, θ^*. The strong convexity of ℓ_n implies that

$$\frac{\eta}{2}\|\widehat{\Delta}\|_2^2 \leq \ell_n(\theta^* + \widehat{\Delta}) - \ell_n(\theta^*) - \langle \nabla \ell_n(\theta^*), \widehat{\Delta} \rangle. \tag{15}$$

Since $\widehat{\theta}$ is the true minimizer of the ℓ_n, we observe that $\ell_n(\theta^* + \widehat{\Delta}) - \ell_n(\theta^*) \leq 0$. Together with (15), this implies that

$$\frac{\eta}{2}\|\widehat{\Delta}\|_2^2 \leq -\langle \nabla \ell_n(\theta^*), \widehat{\Delta} \rangle \leq \|\nabla \ell_n(\theta^*)\|_\infty \|\widehat{\Delta}\|_1 \leq \sqrt{d}\|\nabla \ell_n(\theta^*)\|_\infty \|\widehat{\Delta}\|_2,$$

which further implies that

$$\|\widehat{\Delta}\|_2 \leq \frac{2\sqrt{d}}{\eta}\|\nabla \ell_n(\theta^*)\|_\infty. \tag{16}$$

Therefore, once we establish the bound on $\|\nabla \ell_n(\theta^*)\|_\infty$, we obtain an upper bound on $\|\widehat{\Delta}\|_2 = \|\widehat{\theta} - \theta^*\|_2$. We begin by showing that $\|\nabla \ell_n(\theta^*)\|_\infty$ can be upper bounded with high probability in both cases.

LEMMA 1. *Under the previous assumptions, in the deterministic design setting, we have with probability at least* $1 - \frac{1}{n}$,

$$\|\nabla \ell_n(\theta^*)\|_\infty \leq cB\sqrt{\frac{\log(nd)}{n}}. \tag{17}$$

For the randomized design, we have with probability at least $1 - \frac{1}{n}$,

$$\|\nabla \ell_n(\theta^*)\|_\infty \leq c\psi\sqrt{\frac{\log(nd)}{n}}. \tag{18}$$

W e note the j-th component of $\nabla \ell_n(\theta^*)$ takes the following form,

$$[\nabla \ell_n(\theta^*)]_j = \frac{1}{n}\sum_{i=1}^{n} W_{ij},$$

where $W_{ij} = \left(\frac{e^{\langle x_i, \theta^* \rangle}}{1+e^{\langle x_i, \theta^* \rangle}} - y_i\right) x_{ij}$ Conditioned on x_i, W_{ij} is a zero-mean bounded random variable with $|W_{ij}| \leq |x_{ij}| \leq B$. For the fixed design setting, applying Hoeffding's inequality,

$$\Pr\left(|[\nabla \ell_n(\theta^*)]_j| \geq t\right) \leq 2\exp\left(-\frac{nt^2}{2B^2}\right).$$

By a union bound, we have

$$\Pr\left(\|\nabla \ell_n(\theta^*)\|_\infty \geq t\right) \leq 2\exp\left(\log(d) - \frac{nt^2}{2B^2}\right). \tag{19}$$

By setting $t = B\sqrt{\frac{2\log(2nd)}{n}}$, we make the R.H.S. of (19) equal to $\frac{1}{n}$, which gives the result in (17).

In the randomized design setting, W_{ij} is a centered sub-Gaussian random variable with the norm bounded above by ψ. To see this, observe that

$$\mathbb{E}\exp\left(tW_{ij}\right) \leq \mathbb{E}\exp\left(|tW_{ij}|\right) \leq \mathbb{E}\exp\left(|tx_{ij}|\right) \leq \mathbb{E}\exp\left(|t|\langle \operatorname{sign}(x_{ij})e_j, x_i\rangle\right) \leq \exp\left(ct^2\psi^2\right)$$

for some constant c. Applying Hoeffding's inequality this implies that,

$$\Pr\left(|[\nabla \ell_n(\theta^*)]_j| \geq t\right) \leq 2\exp\left(-\frac{c_1 nt^2}{\psi^2}\right).$$

By a union bound, we then have that

$$\Pr\left(\|\nabla \ell_n(\theta^*)\|_\infty \geq t\right) \leq 2\exp\left(\log(d) - \frac{c_1 nt^2}{\psi^2}\right). \tag{20}$$

By setting $t = \psi\sqrt{\frac{\log(2nd)}{c_1 n}}$, we make the R.H.S. of (20) equal to $1 - \frac{1}{n}$, which gives the result in (18).

In the next lemma, we identify the required strong-convexity parameter η of ℓ_n.

LEMMA 2. *Under the previous assumptions, for deterministic design setting, we have that ℓ_n is strongly convex with*

$$\eta = \frac{\exp(RB)}{4(1+\exp(RB))^2} \cdot \rho. \tag{21}$$

In the randomized design setting, as long as $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$ for some constant $C_{cp}(\psi)$ only depending on ψ, ℓ_n is strongly convex with

$$\eta = \frac{\exp(RB)}{4(1+\exp(RB))^2} \cdot \rho, \tag{22}$$

with probability at least $1-2(\frac{1}{n})^d$.

T he Taylor expansion of ℓ_n implies that for some $\alpha \in (0,1)$ we have

$$\ell_n(\theta^* + \widehat{\Delta}) - \ell_n(\theta^*) - \langle \nabla \ell_n(\theta^*), \widehat{\Delta} \rangle = \frac{1}{2n} \sum_{i=1}^{n} \frac{\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta})}{(1+\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta}))^2} \widehat{\Delta}^T \left(x_i x_i^T\right) \widehat{\Delta}. \tag{23}$$

Further,

$$|x_i^T\theta^* + \alpha x_i^T \widehat{\Delta}| \leq \|x_i\|_\infty \|\theta^* + \alpha\widehat{\Delta}\|_1 \leq \|x_i\|_\infty \left(\alpha\|\theta^*\|_1 + (1-\alpha)\|\widehat{\theta}\|_1\right) \leq RB.$$

Since the function $\frac{\exp(a)}{(1+\exp(a))^2} \leq \frac{1}{4}$ is an even function and monotonically decreasing as $|a|$ increases, we have

$$\frac{\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta})}{(1+\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta}))^2} \geq \frac{\exp(RB)}{(1+\exp(RB))^2}, \tag{24}$$

which further implies that,

$$\begin{aligned} \frac{1}{2n} \sum_{i=1}^{n} \frac{\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta})}{(1+\exp(x_i^T\theta^* + \alpha x_i^T \widehat{\Delta}))^2} \widehat{\Delta}^T \left(x_i x_i^T\right) \widehat{\Delta} &\geq \frac{1}{2} \frac{\exp(RB)}{(1+\exp(RB))^2} \widehat{\Delta}^T \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T\right) \widehat{\Delta} \\ &\geq \frac{1}{2} \frac{\exp(RB)}{(1+\exp(RB))^2} \lambda_{\min}(\Sigma_n) \|\widehat{\Delta}\|_2^2. \end{aligned}$$

In the deterministic setting, by the assumption that $\lambda_{\min}(\Sigma_n) \geq \frac{\rho}{2}$, we obtain the desired result in (21).

In the randomized setting, by Corollary 5.50 from (Vershynin 2012), for any $\epsilon \in (0,1)$ and $t \geq 1$, when $n \geq C_{cp}(\psi)\left(\frac{t}{\epsilon}\right)^2 d$, where $C_{cp}(\psi)$ is a constant depends only on ψ, with probability at least $1 - 2\exp(-t^2 d)$,

$$\|\Sigma_n - \Sigma\|_{\text{op}} \leq \epsilon.$$

By setting $\epsilon = \frac{1}{2}\min(\rho, 1)$ and $t = \sqrt{\log(n)}$, we have with probability at least $1 - 2(\frac{1}{n})^d$,

$$\|\Sigma_n - \Sigma\|_{\text{op}} \le \frac{1}{2}\rho,$$

provided that $n \ge \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$. By Weyl's theorem, we have

$$|\lambda_{\min}(\Sigma_n) - \lambda_{\min}(\Sigma)| \le \|\Sigma_n - \Sigma\|_{\text{op}} \le \frac{1}{2}\rho,$$

which further implies that,

$$\lambda_{\min}(\Sigma_n) \ge \lambda_{\min}(\Sigma) - \frac{1}{2}\rho \ge \frac{1}{2}\rho.$$

This completes the proof.

Combining Lemma 1 and 2 with (16), we obtain the following theorem by following the steps sketched above.

THEOREM 1. *Under the previous assumptions, in deterministic design setting, with probability at least* $1 - \frac{1}{n}$,

$$\|\widehat{\theta} - \theta^*\|_2 \le c\frac{B}{\rho}\frac{(1+\exp(RB))^2}{\exp(RB)}\sqrt{\frac{d\log(nd)}{n}}. \tag{25}$$

In the randomized design setting, as long as $n \ge \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$ *for some constant* $C_{cp}(\psi)$ *only depending on* ψ, *w.p. at least* $1 - \frac{1}{n} - 2(\frac{1}{n})^d$,

$$\|\widehat{\theta} - \theta^*\|_2 \le c\frac{\psi}{\rho}\frac{(1+\exp(RB))^2}{\exp(RB)}\sqrt{\frac{d\log(nd)}{n}}, \tag{26}$$

where c *is a universal constant.*

Given that (14) holds with high probability, we now present the associated bound on expected revenue. In what follows, we fix a feature vector z. Define the optimized prices

$$\hat{p} := \max_{k\in[K]} f_z(p, \widehat{\beta}, \widehat{\gamma}) \quad \text{and} \quad p^* := \max_{k\in[K]} f_z(p, \beta^*, \gamma^*).$$

We now provide a bound on the loss in expected revenue for choosing to offer $\hat{p}$ over the unknown p^* and demonstrate that this revenue loss decreases at a quantifiable rate as the sample size n is increased.

Theorem 2. *With high probability, as long as* $n \geq \frac{4C_{cp}(\psi)\log(n)d}{\min(\rho,1)^2}$*, the loss in expected revenue associated with choosing* $\hat{p}$ *over* p^* *can be bounded as follows:*

$$f_z(p^*,\beta^*,\gamma^*) - f_z(\hat{p},\beta^*,\gamma^*) \leq \left(\max_{k\in[K]} p^k\right) \frac{C'_{cp}(R,B,\psi)}{\rho} \sqrt{\frac{d^2\log(nd)}{n}},$$

for all bounded feature vectors z *where* $C'_{cp}(R,B,\psi)$ *is a constant only depending on* R, B *and* ψ*.*

F or a fixed k, we have

$$\begin{aligned}
\left| f_z(p^k,\beta^*,\gamma^*) - f_z(p^k,\widehat{\beta},\widehat{\gamma}) \right| &= \left| \frac{p^k}{1+\exp\left(-\left(\beta^*_k + \sum_j \gamma^*_j z_j\right)\right)} - \frac{p^k}{1+\exp\left(-\left(\widehat{\beta}_k + \sum_j \widehat{\gamma}_j z_j\right)\right)} \right| \\
&\leq \frac{p^k}{4} \left| \langle (1,z), (\beta^*_k - \widehat{\beta}_k, \gamma^* - \widehat{\gamma}) \rangle \right| \qquad (27) \\
&\leq \frac{p^k}{4} \|(1,z)\|_2 \left\| (\beta^*_k - \widehat{\beta}_k, \gamma^* - \widehat{\gamma}) \right\|_2 \\
&\leq \left(\max_{k\in[K]} p^k\right) \frac{\sqrt{mB^2+1}}{4} \left\| (\beta^* - \widehat{\beta}, \gamma^* - \widehat{\gamma}) \right\|_2,
\end{aligned}$$

where (27) follows from the fact that the derivative of the function $(1+\exp(-a))^{-1}$ is bounded by $\frac{1}{4}$ for any a. Thus, using the fact that $f_z(\widehat{p},\widehat{\beta},\widehat{\gamma}) \geq f_z(p^*,\widehat{\beta},\widehat{\gamma})$,

$$\begin{aligned}
f_z(p^*,\beta^*,\gamma^*) - f_z(\widehat{p},\beta^*,\gamma^*) &= f_z(p^*,\beta^*,\gamma^*) - f_z(\widehat{p},\widehat{\beta},\widehat{\gamma}) + f_z(\widehat{p},\widehat{\beta},\widehat{\gamma}) - f_z(\widehat{p},\beta^*,\gamma^*) \\
&\leq f_z(p^*,\beta^*,\gamma^*) - f_z(p^*,\widehat{\beta},\widehat{\gamma}) + f_z(\widehat{p},\widehat{\beta},\widehat{\gamma}) - f_z(\widehat{p},\beta^*,\gamma^*) \\
&\leq \left| f_z(p^*,\beta^*,\gamma^*) - f_z(p^*,\widehat{\beta},\widehat{\gamma}) \right| + \left| f_z(\widehat{p},\widehat{\beta},\widehat{\gamma}) - f_z(\widehat{p},\beta^*,\gamma^*) \right| \\
&\leq \left(\max_{k\in[K]} p^k\right) \frac{\sqrt{mB^2+1}}{4} \left\| (\beta^* - \widehat{\beta}, \gamma^* - \widehat{\gamma}) \right\|_2. \qquad (28)
\end{aligned}$$

We can then apply (14) to get the desired result.

As desired, we have bounded the optimality gap of the reward generated by the recommended price $\widehat{p}$ as compared to the optimal price p^*. This bound decreases as $O\left(\frac{1}{\sqrt{(n)}}\right)$, up to logarithmic terms. It is interesting to note that because of the discrete nature of the problem of choosing a price to offer, for large enough n, the bound will guarantee a revenue gap of zero.

5.2. Personalized Assortment Optimization Bounds

We can give a similar bound for the personalized assortment optimization revenue function, with the addition of a factor of J^4, where J is the number of products. The key assumption we make in addition to the assumptions for customized pricing is that the seller has *sufficiently explored* every product. In particular, for each product $j \in \{1, \ldots, J\}$, let

$$\mathcal{I}_j = \{i : j \in S_i\}. \tag{29}$$

We assume that $n_j = |\mathcal{I}_j| > \nu n$ for all j where $\nu \in (0, 1]$ is a constant.

As in the customized pricing model, we assume that the outputs $y_i \in \{0, 1\}$ are independent given each z_i. We also assume bounded inputs, i.e., there exists a constant $B > 0$, such that for any $i \in [n]$ and $k \in [d]$

$$|z_{ik}| \leq B.$$

We again consider both the deterministic design and random design and make similar assumptions concerning the postive definiteness of the empirical covariance matrices in the deterministic case and the true covariance matrix in the randomized setting. In the randomized setting we again assume that the feature vectors follow a sub-Gaussian distribution.

1. For deterministic design, we assume that

   $$\lambda_{\min}(\Sigma_{n_j}) \geq \frac{\rho_j}{2} > 0,$$

   where $\Sigma_{n_j} = \frac{1}{n_j} \sum_{i=1}^{n_j} z_i z_i^T$. Denote $\rho = \min_j \{\rho_j\}$.
2. For the random design, we assume that $\{z_i\}_{i=1}^n$ are i.i.d. sub-Gaussian random vectors with sub-Gaussian norm ψ, i.e.,

   $$\psi = \sup_{\|w\|_2 \leq 1} \|\langle z, w \rangle\|_{\psi_2},$$

   where z has the same distribution as any z_i.

   Further, we assume that

   $$\lambda_{\min}(\Sigma) > \rho > 0, \tag{30}$$

   where $\Sigma = \mathbb{E}(zz^T)$.

Using these assumptions and a proof similar to that used in the case of customized pricing, we derive the following results:

THEOREM 3. *(MLE Parameter Bound for multinomial logit) Under the previous assumptions, for deterministic design, with probability at least* $1-J/n$*,*

$$\|\widehat{\theta}_j - \theta_j^*\|_2 \le c\frac{B}{\rho}\frac{[1+\exp(RB)]^2}{\exp(-RB)}J^2\sqrt{\frac{d\log(nd)}{n\nu}}$$

simultaneously for all products $j=1,\ldots,J$.

For the randomized design, as long as $n \ge \frac{4C_{ao}(\psi)\log(n)d}{\nu\min(\rho,1)^2}$ *for some constant* $C_{ao}(\psi)$ *only depending on* ψ*, with probability at least* $1-J/n-2J(1/n)^d$*,*

$$\|\widehat{\theta}_j - \theta_j^*\|_2 \le c\frac{\psi}{\rho}\frac{[1+\exp(RB)]^2}{\exp(-RB)}J^2\sqrt{\frac{d\log(nd)}{n\nu}}$$

simultaneously for all products $j=1,\ldots,J$.

THEOREM 4. *(Optimality Gap Bound for Personalized Assortment) Under the previous assumptions, with high probability the optimality gap of the algorithm as specified above can be bounded as follows:*

$$f_z(S^*,\theta^*) - f_z(\widehat{S},\theta^*) \le \frac{2C'_{ao}(R,B,\psi)}{\rho}r_1J^4\sqrt{\frac{d^2\log(nd)}{n\nu}}.$$

The proof for Theorem 3 uses the exact method for proving Theorem 1, viewing the assortment optimization negative log-likelihood as a function only of θ_j and fixing all other parameters to their predicted values. Theorem 4 is a straight-forward extension of 2.

6. Extensions of Theory

The theory presented so far been developed assuming our model has been well-specified and implicitly assuming a low-dimensional feature space due to the appearance of the dimension in our revenue bounds. In this section, we demonstrate that it is possible for both of these assumptions to be relaxed.

6.1. Misspecified Model Setting

Sometimes it is unreasonable to assume that the logit model correctly specifies the true underlying probabilities. In this case, the same bounds still hold for a different optimality gap.

Suppose the data $\mathcal{T} = \{(z_1, a_1, y_1), \ldots (z_n, a_n, y_n)\}$ is generated from some underlying random process. We do not consider fixed design here. As a new benchmark, we redefine

$$\theta^* = \arg\min_{\theta} \mathbb{E}(\ell_n(\theta)),$$

so that in this section, θ^* represents the estimated parameters given an infinite amount of i.i.d. data rather than the true parameters. We will refer to θ^* in this section as the oracle estimator.

As before, we assume that we have a number R such that $\|\theta^*\|_1 \leq R$. We assume the inputs z_i to be i.i.d. sub-Gaussian random variables with sub-Gaussian norm ψ, and that the minimum eigenvalue of the covariance matrix of z is at least ρ, where z has the same distribution as each z_i.

All of the results of the previous section will still hold with the new definition of θ^*, provided that

$$\mathbb{E}\nabla\ell_n(\theta^*) = 0. \tag{31}$$

To see that (31) indeed does hold, note that because θ^* is defined as an unconstrained minimum, $\nabla\mathbb{E}(\ell_n(\theta^*)) = 0$. Since each component of $\nabla\ell_n(\theta)$ is integrable, we are justified in switching the order of differentiation and integration, which means that (31) holds. Thus, we can apply the proofs from our previous analysis directly to the misspecified setting. We state just one result, our bound for customized pricing, as an example.

Theorem 5. *With high probability, as long as $n \geq \frac{4C_\psi \log(n) d}{\min(\rho,1)^2}$, the gap between oracle estimator revenue and the revenue achieved with $(\widehat{\beta}, \widehat{\gamma})$ can be bounded as follows:*

$$f_z(p^*, \beta^*, \gamma^*) - f_z(\hat{p}, \beta^*, \gamma^*) \leq \left(\max_{k\in[K]} p^k\right) \frac{C'_{cp}(R, B, \psi)}{\rho} \sqrt{\frac{d^2 \log(nd)}{n}},$$

for all bounded feature vectors z where $C'_{cp}(R, B, \psi)$ is the same constant from Theorem 2.

6.2. High-Dimensional Setting

In previous sections, our bounds are increasing functions of the number of features d. This is reasonable when d remains fixed and n grows large. However, as companies continue to collect more granular and larger quantities of data concerning their customers, there are applications where the number of customer features matches or even exceeds the number of data points. Examples of this could include highly granular GPS information or detailed lists of preference comparisons featuring irrational cycles that are not easily specified without introducing a combinatorial number of parameters.

In such a high-dimensional setting where d is comparable to or larger than n, a bound of the form presented above is of no use. In order to derive analogous bounds for the performance of our method in this setting, we must make assumptions about the sparsity of the true parameter θ^*. Specifically, for some index set $S \subset \{1 \dots d\}$ with $|S| = s \ll d$, we assume that $\theta^* \in \Theta(S) := \{\theta \in \mathbb{R}^d : \theta_j = 0, \ \forall j \notin S\}$. In the well-specified case this means that a small portion of the available feature data can be used to precisely quantify the likelihood of purchase. A natural way to obtain an estimator of such a sparse vector is to use ℓ_1-norm regularization as we have considered in the low-dimensional case.

The sparsity assumption allows us to prove theoretical revenue bounds analogous to those in the low dimensional setting, with dimension parameter d replaced by sparsity parameter s and only a logarithmic dependence on d. We include here the high-dimensional version of our revenue bound for the well-specified random design customized pricing problem.

Using a soft constraint form of the regularization problem, we obtain an estimator given by

$$\widehat{\theta}_{\lambda_n} \in \arg\min_{\theta \in \mathbb{R}^d} \{\ell_n(\mathcal{T}, \theta) + \lambda_n \|\theta\|_1\}. \tag{32}$$

THEOREM 6. *As long as $n > 16(\frac{\kappa_2}{\kappa_1})^2 s \log d$, the objective function error associated with $(\widehat{\beta}, \widehat{\gamma})$, the solution to (32) with regularization parameter $\lambda_n = 4\psi\sqrt{\frac{\log p}{2n}}$, can be bounded as follows:*

$$f_z(p^*, \theta^*) - f_z(\widehat{p}, \theta^*) \leq \left(\max_{k \in [K]} p^k\right) \frac{12B\psi\sqrt{\log d}}{\kappa_1\sqrt{2n} - 4\kappa_2\sqrt{2s\log d}},$$

for all bounded feature vectors z with probability at least $1 - 2d\exp(-\frac{n}{4}) - c_1\exp(-c_2 n) - \frac{2}{d}$. Here κ_1, κ_2, c_1, and c_2 are positive constants depending only on B, ψ, ρ, and Σ.

The proofs are highly technical and follow from Negahban et al. (2010) and Negahban et al. (2012). The rest of our results can be extended to the high-dimensional setting using similar techniques.

7. Experiments

The theory we have presented so far suggests that our method provides an effective technique for estimating and optimizing data-driven decisions. However, many of the bounds we present are a worst case analysis and we would expect that, on average, the performance of the algorithms we suggest would far exceed these worst case scenarios. To test this intuition, in this section, we describe the results of the experiments on real and simulated data.

7.1. Customized Pricing for Airline Priority Seating

In addition to our theoretical results, we were also fortunate to be able to test our method for data-driven customized pricing with sales data from a European airline carrier. The data set concerns sales of passenger seating reservations, in which, for an extra fee, passengers of this carrier may pre-select a seat on their airplane that will be reserved for them on the day of their flight. The charges for these reservations are small in comparison to the price of the ticket, but optimizing these prices would allow the airline to generate extra revenue at essentially no marginal cost. Due to the potential legal issues with offering prices to customers based on their personal attributes, the data consists solely of information concerning the flight such as date or destination and transaction information such as date and time of website access that could also be used to fairly adjust prices for all customers. This data set is unique in that the carrier was willing to perform random price experimentation and that we observe decisions by customers both to purchase and to reject the reserved seating option.

This data set was collected for flights within a single country over the month of December 2014. During this period, some fraction of all customers who purchased domestic airline tickets was offered the opportunity to purchase a seating reservation at a price randomly selected with equal probability from four candidate prices. We will refer to these four prices in order with Price 1 being the smallest and Price 4 being the highest. The resulting data set consists of around 300,000 transaction records with the treatment price offered, the resulting purchase decision, and data concerning attributes of the flight and of the transaction. Such attributes include the date and time of booking, the date and time of the flight and the corresponding return flight, the origin and destination pair, the number of passengers, and the average ticket price paid among other variables.

The performance metric of interest in this case is the expected revenue from reserved seating per passenger. To allow us to fairly evaluate the success of our method, we began by splitting our data into a training and testing set chronologically, with the earliest 60% of transaction records used for training and the latest 40% used for testing, corresponding to a testing training split on December 18, 2014. This ensures that any method selected based on the training set could have been implemented by the company during the subsequent testing period. The benchmark to which we will compare our method is the performance of the best single price that could have been selected by the company, based on the training data, on the testing set data. Since treatment prices were offered independently at random, we can estimate in an unbiased fashion such a single price performance, by restricting our attention to customers who were offered each specific price during the testing period. The highest single-price expected revenue per passenger on the training set is achieved using the highest price, Price 4. Therefore we will report the test set performance of our method as ratios to the test set performance of this price.

Since each customer was only offered a single price, we are unable to perform a true counterfactual analysis to evaluate algorithm performance when it chooses to offer a price to customer that is different than the one they were offered in the data set. However since treatment prices were offered to customers independently at random, to estimate expected revenue per passenger it is sufficient to restrict our attention to the subset of records for which the best price selected by the algorithm matches the price offered in actuality. Although this limits the size of the test set, it allows us to obtain an unbiased estimate of expected revenue per passenger without the need for potentially erroneous counterfactual assumptions.

In accordance with our algorithm given in section 3, we fit a logistic regression model to predict purchase probability using the treatment price offered as a feature as well as the flight and transaction data discussed previously. We also found it useful to include interaction effects between the treatment price offered and other variables in the data set. To perform feature selection we used a 4-fold cross-validation procedure with our cross-validation metric being the expected revenue per passenger, estimated as specified in the previous paragraph, on the held out subset of data. Our final model was selected as the linear model with the subset of features generating the best performance in the cross-validation trials.

The best models that we identified through this cross validation procedure all considered the effects of the treatment price offered, the day of week of booking (BDW), and the day of week of the return flight if applicable (RFDW). Importantly, they all also used both the interaction between the price offered and BDW and the interaction between the price offered and RFDW.

		Day of Week							
		Sun.	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	NA
BDW	Purchase Likelihood	-				+	+	+	
	Price Sensitivity		+	-		+	-	+	
RFDW	Purchase Likelihood						+	+	-
	Price Sensitivity		+	+	-	-		+	+

Table 1 Displays the effect of booking day of week (BDW) and return flight day of week (RFDW) on seating reservation purchase probability and price sensitivity. Higher purchase likelihood indicates a higher chance to buy at any offered price, while higher price sensitivity indicates more resistance to higher prices compared to lower ones. '+' denotes a relatively higher value than average while '-' denotes a relatively lower value.

Due to the interpretability of logistic regression we are able to draw some conclusions from our models on how the factors above affect willingness to purchase a reserved seating option. These conclusions for BDW and RFDW are summarized in table 2. For example, at any offered price a customer with a return flight scheduled for Friday or Saturday has a higher likelihood of purchase while a comparable customer with no return flight scheduled at all is much less likely to purchase. From the interaction effects we observe that price sensitivity increases for samples when the BDW is Monday, Thursday, and Saturday and is lower when tickets are purchased on Tuesday and Friday. Many of these effects align with the traditional division between leisure travelers and business travelers, who are assumed to be less price sensitive.

Using the best model selected by the process above, our price selection technique generated an increase in estimated expected revenue per passenger of 7.49% over the performance of statically offering Price 4. To demonstrate the robustness of our technique we also computed the performance of the top ten models we identified using revenue performance on the cross-validation and note that all of these models have similar performance to the final model. These estimated expected revenue performance on the test set of our strategies and each of the single price strategies are displayed in Figure 1.

It is interesting to note that although Price 4 generated the highest expected revenue on the training set, this was not the case on testing set. For those dates later in the month, Price 3 happened to yield the best performance generating 2.95% more revenue per passenger than the static Price 4. This highlights the non-stationarity of real transaction data and motivates the need to split the training and testing set chronologically rather than at random for fair evaluation.

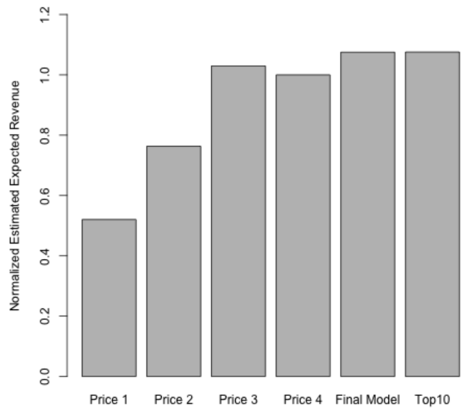

Figure 1 **Performance of pricing strategies over offering Price 4. The first four bars represent the performance of offering single prices. Final Model plots the performance of the best cross-validated model. Top 10 plots the average performance of the 10 best models identified in the cross-validation process.**

7.2. Numerical Experiments

In addition to the airline data test, we performed simulations for both customized pricing and personalized assortment optimization. We tested our algorithm against an aggregate approach that does not use feature information and against an oracle who knows the true distribution.

7.2.1. Well-specified Model For customized pricing we defined a problem class by specifying each of the following:

- Number of prices $K \in \{2,4,10\}$
- Number of features $d \in \{5,10,15\}$
- Amount of training data $n \in \{100,300,500\}$.

For each problem class we performed $N = 100$ trials. For each trial, we generated a price set of size K by random sampling from a uniform distribution on $[5,20]$, with the lowest price anchored at 5. We also generated a d-dimensional true parameter vector γ^*, where each dimension is chosen i.i.d from a normal distribution with mean zero and standard deviation 1.5, and a K-dimensional parameter vector β as the order statistics of K i.i.d. normal random variables with mean zero and standard deviation 2.5 sorting from lowest to highest. The difference in standard deviations for the two parameters allows the price effect to dominate the effects of the other features. The sorting of the β_k is motivated by the fact that in almost all cases, demand for a product is decreasing in its price. We then generated n data points as a training set, where each data point consists of a feature vector z of size d, drawn i.i.d. from a multivariate normal distribution, a price drawn uniformly at random from the constructed set, and a purchase decision given according to the logistic regression model with the true parameters γ^* and β^*. The distribution of the vectors z had mean zero and a covariance structure such that $\mathrm{var}(Z_i) = 1$ and $\mathrm{cov}(Z_i, Z_j) = 0.3$, $i \neq j$. We trained the following using the training data:

1. The Personalized Revenue Maximization Algorithm (PRM),
2. The I-PRM Algorithm (a variant of PRM described below), and
3. A single-price policy.

In cases with a small number of data samples, we would expect information about the monotonicity of the β_k parameters to be helpful. Thus, the I-PRM algorithm (for Isotonic-constraint PRM) performs the same maximum likelihood estimation as PRM, but with the added constraint that β must lie in the cone $\{x \in \mathbb{R}^K \mid x_1 \leq \cdots \leq x_K\}$.

We generated a test set of 1000 data points consisting of feature vectors and a reservation price generated according to the true model. We then used the test set to calculate empirical expected revenue for each method.

Figure 2 shows the performance of all three methods for $K = 10$ under various dimensions of the feature space. The isotonic-constrained version of PRM Algorithm slightly outperforms the regular version, and both are significantly better than a single price strategy. Not very many samples are required for the algorithms to recover almost all of the oracle revenue.

We found that the PRM Algorithm consistently outperforms the single price policy by 4-20% on average for each problem class, with the performance improvement becoming more pronounced as K and d increase. In the case of 2 prices, the algorithm recovers at least 94% of the oracle revenue with only 100 training samples, and closes to 1% with 300 samples.

8. Extensions and Future Work

Beyond problems in revenue management our approach is relevant in many other situations in which decisions resulting in discrete outcomes can benefit from taking into account explicit contextual information. One such example is in online advertisement allocation in which we would like to predict click-through rates and make the optimal advertisement

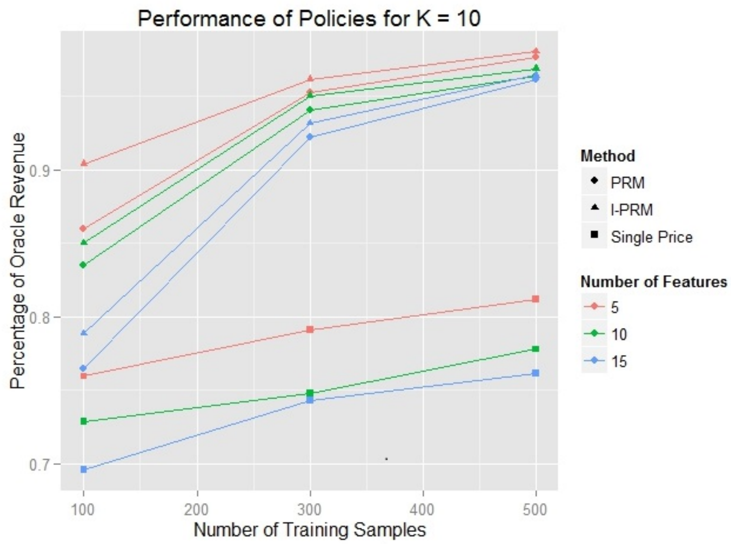

Figure 2 **Performance of three pricing methods in the well-specified case, with $K = 10$ prices**

selection taking into account information we have about each viewer. Another example application is crowdsourcing in which we would like to specialize our work schedule based on information we have gathered concerning our workers, the available tasks, and the interaction between their attributes. Finally, beyond the specific domain of operations management we envision applications in personalized medicine in which the likelihood of success of a treatment or the probability of disease could be predicted and decisions optimized by taking into account information concerning each patient.

Although our framework is quite general, we make some simplifying assumptions in our analysis that could be relaxed in future work. Many of the papers that study dynamic pricing and assortment optimization do so with inventory constraints over multiple periods, whereas our results do not explicitly consider inventory and are inherently myopic due to the offline nature of logistic regression. Due to this offline nature, our method does not work to dynamically learn and optimize over time, and it would be interesting to examine how these results could be adapted to such a setting. An alternation scheme between exploration and exploitation phases such as that suggested by Rusmevichientong et al. (2010) would allow for asymptotic convergence, but the logistic regression would need to be recomputed at step, which could be cumbersome even with a warm-start as the size of the data set grows large. A potential remedy for this would be to learn from a finite window of past data. Such a scheme could account for changes in parameters over time in which case asymptotic convergence becomes meaningless.

Finally, assuming a logistic model is only one way to gain the theoretical tractability necessary to produce finite-sample performance guarantees. There are likely to be many other types of models which are capable of producing similar results. We believe that future work in broadening the scope of this style of analysis could be fruitful.

References

Xavier Amatriain. Big & personal: data and models behind netflix recommendations. In *Proceedings of the 2nd International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications*, pages 1–6. ACM, 2013. 1

Yossi Aviv, Gustavo Vulcano, Ö Özer, and R Phillips. Dynamic list pricing. *The Oxford Handbook of Pricing Management. Oxford University Press, Oxford*, pages 522–584, 2012. 4

Goker Aydin and Serhan Ziya. Technical note-personalized dynamic pricing of limited inventories. *Operations Research*, 57(6):1523–1531, 2009. 5

P. P. Belobaba. Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, 37:183–197, 1989. 5

Fernando Bernstein, A. Gürhan Kök, and Lei Xie. Dynamic assortment customization with limited inventories. Working Paper, 2011. 6

Josef Broder and Paat Rusmevichientong. Dynamic pricing under a general parametric choice model. *Operations Research*, 60(4):965–980, 2012. 4

Peter Bühlmann and Sara van de Geer. *Statistics for High-Dimensional Data: Methods, Theory, and Applications*. Springer-Verlag, 2011. 15

Alexandre X. Carvalho and Martin L. Puterman. Learnind and pricing in an internet environment with binomial demands. *Journal of Revenue and Pricing Management*, 3(4):320–336, 2005. 3, 5

C-C Henry Chan, Chi-Bin Cheng, and Wen-Chen Hsien. Pricing and promotion strategies of an online shop based on customer segmentation and multiple objective decision making. *Expert Systems with Applications*, 38(12):14585–14591, 2011. 1

N. Golrezaei, H. Nazerzadeh, and P. Rusmevichientong. Real-time optimization of personalized assortments. *Management Science*, 60(6):1532–1551, 2014. 3, 6

N. Bora Keskin and Assaf Zeevi. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research*, 62(5):1142–1167, 2014.

A.G. Kök, M.L. Fisher, and R. Vaidyanathan. *Retail Supply Chain Management*, chapter Assortment Planning: Review of Literature and Industry Practice, pages 1–55. 2008. 5

B. Letham, W. Sun, and A. Sheopuri. Latent variable copula inference for bundle pricing from retail transaction data. *Proceedings of the 31st International Conference on Machine Learning*, 2014.

Greg Linden, Brent Smith, and Jeremy York. Amazon. com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE*, 7(1):76–80, 2003. 1

BPS Murthi and Sumit Sarkar. The role of the management sciences in research on personalization. *Management Science*, 49(10):1344–1362, 2003. 1

Sahand N. Negahban, Pradeep Ravikumar, Martin J. Wainwright, and Bin Yu. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. *arXiv pre-print*, 2010. arXiv:1010.2731v1. 23

Sahand N. Negahban, Pradeep Ravikumar, Martin J. Wainwright, and Bin Yu. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. *Statistical Science*, 27(4): 538–557, 2012. 23

Serguei Netessine, Sergei Savin, and Wenqiang Xiao. Revenue management through dynamic cross selling in e-commerce retailing. *Operations Research*, 54(5):893–913, 2006. 5

Cynthia Rudin and Gah-Yi Vahn. The big data newsvendor: Practical insights from machine learning. working paper, 2014. 4

P. Rusmevichientong, Z.J. Max Shen, and D. B. Shmoys. Dynamic assortment optimization with a multinomial logit choice model and a capacity constraint. *Operations Research*, 58(6):1666–1680, 2010. 6, 13, 29

P. Rusmevichientong, D.B. Shmoys, C. Tong, and H. Topaloglu. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11): 2023–2039, 2014. 6, 13

Kalyan Talluri and Garrett van Ryzin. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):15–33, 2004a. 5, 13

Kalyan T. Talluri and Garrett J. van Ryzin. *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, 2004b. 1

Garrett van Ryzin and S Mahajan. On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, 45(11):1496–1509, 1999. 5

Roman Vershynin. *Compressed Sensing, Theory and Application (*, chapter 5, pages 210–268. Cambridge University Press, 2012. 17

Gustavo Vulcano, Garrett van Ryzin, and Richard Ratliff. Estimating primary demand for substitutable products from sales transaction data. *Operations Research*, 60(2):313–334, 2008. 6