WZB

Wissenschaftszentrum Berlin
für Sozialforschung

Qi Zhao, Yongfeng Zhang, Yi Zhang
Daniel Friedman

Recommendation based on multi-product utility maximization

Discussion Paper

SP II 2016–503

June 2016

Research Area
Markets and Choice

Research Professorship
Market Design: Theory and Pragmatics

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Affiliation of the authors:

Qi Zhao, School of Engineering, University of California, Santa Cruz
(qzhao2@ucsc.edu)

Yongfeng Zhang, School of Engineering, University of California, Santa Cruz, and
Department of Computer Science & Technology, Tsinghua University, Beijing
(zhangyf07@gmail.com)

Yi Zhang, School of Engineering, University of California, Santa Cruz (yiz@ucsc.edu)

Daniel Friedman, WZB Berlin Social Science Center, and Economics Department,
University of California, Santa Cruz (dan@ucsc.edu)

ABSTRACT

Recommender systems often recommend several products to a user at the same time, but with little consideration of the relationships among the recommended products. We argue that relationships such as substitutes and complements are crucial, since the utility of one product may depend on whether or not other products are purchased. For example, the utility of a camera lens is much higher if the user has the appropriate camera (complements), and the utility of one camera is lower if the user already has a similar camera (substitutes). In this paper, we propose *multi-product utility maximization* (MPUM) as a general approach to account for product relationships in recommendation systems. MPUM integrates the economic theory of consumer choice theory with personalized recommendation, and explicitly considers product relationships. It describes and predicts utility of product bundles for individual users. Based on MPUM, the system can recommend products by considering what the users already have, or recommend multiple products with maximum joint utility. As the estimated utility has monetary unit, other economic based evaluation metrics such as consumer surplus or total surplus can be incorporated naturally. We evaluate MPUM against several popular baseline recommendation algorithms on two off-line E-commerce datasets. The experimental results showed that MPUM significantly outperformed baseline algorithm under top-K evaluation metric, which suggests that the expected number of accepted/purchased products given K recommendations are higher.

Categories and Subject Descriptors

M.5.4 [**Applied Computing**]: Law, Social and Behavioral Sciences- *Economics*; H.3.3 [**Information Search and Retrieval**]: Information Filtering

Keywords

Recommendation Systems; Utility; Product Portfolio; Computational Economics

1. INTRODUCTION

E-commerce has grown rapidly in recent years and it has become increasingly popular shopping venue. Due to the large number of products and massive information, product recommender systems (RS) help by learning consumer preferences and discovering products a consumer find most valuable among other products. RS has proven to be important for E-commerce websites[28] and are widely adopted in industry. For example, Amazon's "who bought this also bought these" or Target's "guests who viewed this ultimately bought".

It has been well recognized that products are related. Two products could be *substitutes* - buy A instead of B or *complements* - buy A together B. Identifying and making use of such relationships are useful for recommendation systems. For example, knowing a consumer's recent purchase of a digital camera, the recommender system should avoid recommend more digital cameras and instead recommend matching lens or photography books. Yet another example is to recommend a shower faucet and a matching valve at the same time.

In economics, utility is used to measure the value of a product perceived to the consumer and it is fundamental for describing and predicting consumer choices. With a utility metric, a good recommendation should be product(s) with the biggest utility for a given consumer. The existence of inter-product relationship makes modeling product utility non-trivial task. For example, how much utility an additional camera provides and how much a lens provides given the camera? The task becomes even harder when the relationship is less obvious - e.g. what is the utility of the camera given iPhone 6S with a built in camera? To answer these questions, a principled approach is needed to quantitatively measure the total utility of two or more products.

In this paper, we propose to how to measure the total utilities for multiple products and recommend by Multi-product Utility Maximization (MPUM). We first extend the famous Cobb-Douglas utility function by expressing the relationship of two products in an unified manner and derive two utility functional forms that meet our requirements. Then we extend the pairwise utility function to more than two products. We assume user make choices to maximize multi-product utility and use multinomial consumer choice model for that, then the multi-product utility model can be learned to maximize the likelihood of observed user data. MPUM can be easily integrated to derive other economic metrics such as consumer marginal utility for recommendations [32], as the estimated utility by MPUM has monetary unit and can be directly used with product price and production cost.

The rest of the paper is organized as follows: we review the related work in Section 2, and introduce some basic definitions and concepts that form the basis of this work in Section 3. In Section 4 and Section 5 we propose our MPUM framework as well as the personalized transaction-based recommendation strategy, respectively, and further present extensive experimental results based on two different real-world datasets in Section 6. We finally conclusion this work with some of the future research directions in Section 7.

2. RELATED WORK

The advent of the internet resulted in large sets of user data. Consequently automatic recommendation algorithms, such as collaborative filtering algorithms, content-based filtering algorithms, and hybrid algorithms, are becoming popular in online stores. **Collaborative filtering** is based on the assumption that users with similar tastes for previous items will have similar preferences for new items, so the algorithm recommends items ranked highly by users deemed similar to the current user[8, 12, 31, 27, 22, 26]. Such algorithms fall into two main categories. Memory based collaborative filtering algorithms predict the unknown rating for a user on an item based on the weighted aggregation of ratings of other (usually the K most similar) users for the same item. Model based collaborative filtering algorithms use the collection of ratings to fit model parameters, and then make predictions based on the fitted models. These include aspect models, flexible mixture models and factorization models [4, 13, 14, 18]. **Content based filtering** is based on the assumption that the features (meta data, words in description, price, tags, visual features, etc.) used to describe the items that a user likes or dislikes tell much about the user preferences. It usually recommends new items similar to previous items the user liked. The underlying research focuses on estimating a user's profile from explicit feedback on whether she liked previous items. Researchers have tried different methods such as logistic regression, support vector machines, Rocchio, language models, Okapi, and pseudo-relevance feedback. **Hybrid recommendation algorithms** combine collaborative filtering with content based filtering using linear functions or learning to rank methods usually perform better than either filtering method alone.

Most above recommendation methods predict individual product score for each user and rank products accordingly, without considering relationships between the recommended items. One major problem is that the top ranked recommendations might be very similar or even duplicate, which usually is not desirable. To address this issue, researchers proposed to diversify the recommendation results[11, 19, 37, 15]. As the potential benefits of diversity to individual users and business are huge, diversity problem has been heavily studied, mainly on other datasets such as news, movies and music [5][1]. Diversity is used either at recommendation candidates selection, at the item score prediction stage, or at the top-N product re-ranking/filtering stage after individual item scores have been predicted. A typical approach is to introduce certain diversity measure such as the number of categories/singers, relative share of recommendations above or below a certain popularity rank percentile [2], or measure over product graph [1]. Another approach is to use measures that will achieve diversity indirectly, such as the risk of a user portfolio of multiple products [29]. Although diversity is not the focus of this paper, the proposed method will lead to diversity naturally as the result of Diminishing Marginal Utility. How to trade off diversity and accuracy, and how to diversify differently for different product categories will be inferred from the utility model learned from consumer choice data.

The most related work to ours is [16]. This paper focuses on clothes category and assumes products purchased together are complements and products viewed together are substitutes. They showed that the co-viewed and co-purchased products relationships can be discovered based on visual appearances of the cloths. However, the relationships exist for different reasons for different products, and the visual methods won't generalize to all other product categories. In contrast, our approach is very general and directly leads to multi products recommendation results.

Another line of research related to this paper are work about *next basket* recommendation problem, which models the sequential pattern of user purchases and recommend a set of items for user's next visit based on previous purchases. A series of methods have been developed for next basket recommendation [21, 33, 10, 6, 34], among which the Hierarchical Representation Model (HRM) [34] represents state of the art. HRM combines general taste by conventional CF and information from previous transaction aggregated by a nonlinear function. Although this paper is not focused for the next basket recommendation problem, the proposed multi-product utility model can be applied to solve this problem. Assuming products the user have purchased before are already in the set and fixed, the system just needs to find and recommend more products to optimize the total utility for the user.

In recent years, there are some efforts on bringing economics principles into e-commerce recommendation systems. In [32], the authors propose to adopt the law of diminishing marginal utility at individual product level so that perishable and durable products are treatd differently. In [36], a mechanism is developed to estimate consumer's willingness-to-pay (WTP) in E-commerce setting and the estimated WTP is used to price product at individual level so that seller's profit is maximized. In [35], a total surplus based recommendation framework is proposed to match producer and consumer so that the total benefit is maximized. Our research falls into this direction and tries to handle the multi product recommendation problem based on solid economics principles and practical recommendation techniques.

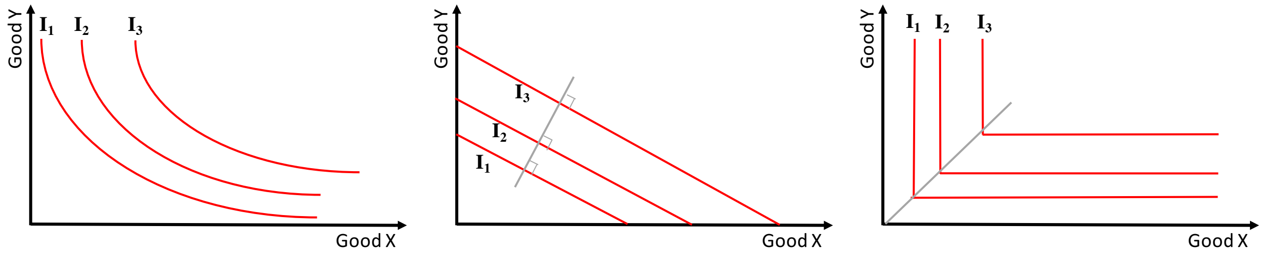
In particular, recognition of product substitutability and complementarity has been considered important for the study the demand of one product affected by other products[3, 24, 25, 30]. Our proposed research is motivated by these existing economics research.

3. BASIC COMPONENTS

In this section, we design some of the key components for our model, and these components will be further integrated into our Multi-Product Utility Maximization (MPUM) framework later.

3.1 Utility

In economics, *utility* is a measure of one's preference over some set of goods or services. It is an important concept that serves as the basis for the rational choice theory [7]. A consumer's total utility for a given set of goods is the



(a) General case (b) *Perfect substitutional* products (c) *Perfect complementary* products

Figure 1: Illustrative indifference curves for common pairs, perfect substitutional pairs, and perfect complementary pairs. The utilities of the three illustrative curves satisfy $I_1 < I_2 < I_3$.

consumer's satisfaction experienced from consuming these goods.

Utility $U(q)$ is usually a function of the consumed *quantity* q. $U(q)$ is inherently governed by the *Law Of Diminishing Marginal Utility* [23], which states that as a person increases the consumption of a product, there is a decline in the marginal utility that the person derives from consuming each additional unit of the product, i.e., $U''(q) < 0$, while the marginal utility keeps positive, i.e., $U'(q) > 0$.

For example, one who is extremely hungry may obtain a huge amount of satisfaction when consuming the first service of bread, but less satisfaction can be obtained when he/she continues to consume when feeling full.

Economists have introduced various functional forms for utility, such as *Cobb-Douglas* utility, *Constant Elasticity of Substitution* (CES) utility, *quasilinear* utility and so on. As each utility function has its own assumptions and limitations, care should be taken when using it, especially for our use cases where more than one products are involved. As we will see later, modeling multi-product utility is essential for us. Instead of simply laying out the function form adopted in the paper, we feel it worth the effort of explaining the motivation behind the choice of the utility function.

3.2 Indifference Curves

In economics, *indifference curves* are used to describe the relationship between arbitrary two products. More specifically, economists are interested in knowing how one product can be substituted by the other product. As illustrated in Figure 1, each indifference curve represents the compositions of two given products that give the same utility. Indifference curves should possesses the following properties,

- Any two points of the same curve give the same utility.
- Curves do not intersect.
- The tangent of any point on the curve is negative, meaning the increment of the consumed quantity of one product requires the decrease of the other product so that the utility keeps unchanged.

Let q_j, q_k denote the consumed quantity of product j and product k respectively. Since points of the same curve have the same utility, the total derivative at any point should be 0, we have:

$$\begin{aligned} \mathrm{d}U(q_j, q_k) &= \frac{\partial U}{\partial q_j}\mathrm{d}q_j + \frac{\partial U}{\partial q_k}\mathrm{d}q_k \\ &= U'_{q_j}\mathrm{d}q_j + U'_{q_k}\mathrm{d}q_k = 0 \end{aligned} \tag{1}$$

Let $h(q_j, q_k) = \frac{\mathrm{d}q_j}{\mathrm{d}q_k}$ denote the Marginal Rate of Substitution (MRS) at point (q_j, q_k), we have,

$$h(q_j, q_k) = \frac{\mathrm{d}q_j}{\mathrm{d}q_k} = -\frac{U'_{q_k}}{U'_{q_j}} \tag{2}$$

Intuitively, The larger $|h(q_j, q_k)|$, the more consumption of product j is need to compensate the decrease of the consumption of product k. As a matter of fact, MRS can fully capture the relationship between two products. To understand this, it might be helpful by looking at the three indifference curve patterns as illustrated in Figure 1. Each pattern represents a typical product relationship. Figure 1a corresponds to the generic case where MRS transits smoothly; Figure 1b corresponds to *perfect substitutes* where the MRS is a constant. In other words, two products can be exchanged at a fixed rate at any time. This can happen when two products are interchangeable, e.g. identical pens except they differ in color and consumer is indifferent in color. Figure 1c corresponds to *perfect complements* where the utility is determined by the minimum of the two product quantity. One might understand this by thinking about the utility of left and right shoes - given a certain quantity of left shoes, the utility will not change by having more right shoes than left shoes and vice versa. Compared to Figure 1a and 1b, the MRS of perfect complements is more tricky: it changes from infinity to zero at certain point. So far, it can be seen that MRS can indicate whether two products are substitutes or complements.

4. MULTIPLE PRODUCT UTILITY MAXIMIZATION FRAMEWORK

In this section, we provide detailed formal treatment of the whole framework by putting the aforementioned essentials together.

4.1 Modeling Marginal Rate of Substitution

First, our goal is to find a proper utility functional form for $U(q_j, q_k)$ so that it can capture all possible products relationships shown in Figure 1. However, the right functional form for the utility function is not obvious, and its not practical for us to try all possible alternatives of $U(q_j, q_k)$ by testing them against the cases in Figure 1 to see how well the utility function can model substitutional and complementary products. Since product substitute and complementary relationships are better illustrated by MRS, we propose to find a proper functional form for MRS, based on which to recover the utility function by solving partial differential equations.

Table 1: Choices of $h(q_k, q_j)$ and its respective utility function. $z(\cdot)$ denotes any monotonic function. Please refer to text for more details.

	Polynomial	Exponential
$h(q_j, q_k)$	$-\frac{a}{1-a}(\frac{q_j}{q_k})^b$ $0 < a < 1, b > 0$	$-\frac{a}{1-a}e^{b(q_j-q_k)}$ $0 < a < 1, b > 0$
$U(q_j, q_k)$	$z\left(aq_j^{1-b} + (1-a)q_k^{1-b}\right)$	$z\left(ae^{-bq_k} + (1-a)e^{-bq_j}\right)$

As MRS can be derived from indifference curve (i.e. $U(q_j, q_k) = const.$), we can alternatively express q_j as a function of q_k, i.e., $q_j = f(q_k)$. The MRS defined in Eq. (2) becomes,

$$\frac{\mathrm{d}q_j}{\mathrm{d}q_k} = f'(q_k) = h(q_j, q_k) \tag{3}$$

where h is the MRS function that we need to decide.

When choosing h, we are mainly concerned about two aspects of h: mathematical convenience and flexibility. Thus we propose to consider two choices listed in Tab 1: polynomial functional form and exponential functional form.

Regardless the specific form of h, the problem of recovering $U(q_j, q_k)$, or $f(q_k)$, boils down to solving the differential function as Eq. 3 for f. Let us see how things pan out for each alternative of h in Tab. 1.

4.1.1 Polynomial Function

We first take a brief look at h to see whether it is expressive enough, i.e., whether it can describe the three cases shown in Figure 1. The answer is positive. When $b = 1$, the resulted MRS is constant $\frac{a}{1-a}$, which is for the case of perfect substitutes; when b goes to very large ($+\infty$), h is large when $\frac{q_j}{q_k} > 1$ and immediately drops to near 0 when $\frac{q_j}{q_k} < 1$, corresponding to the perfect complements case; when $0 < b < 1$, the resulted MRS is for the general case shown in Figure 1a.

After applying some differential equation tricks to Eq. (3), we reach the following equation,

$$\left(aq_j^{1-b} + (1-a)q_k^{1-b}\right) = const. \tag{4}$$

Let's remind ourselves that MRS is defined when utility is set to unknown constant. The above equation suggests that the utility function might be some monotonic function of the left side of the above equation, namely,

$$U(q_j, q_k) = z\left(aq_j^{1-b} + (1-a)q_k^{1-b}\right) \tag{5}$$

where $z()$ is any monotonic function such as log and power. In particular, when $z(x) = x^{\frac{1}{1-b}}$, it results in the well known *Constant Elasticity Substitution* (CES) utility function in Economics, and $s = \frac{1}{b}$ is called the *Elasticity of Substitution*, which denotes the degree of substitutability/complementarity between a pair of products. Specifically, the utility function models (perfect) substitutional product pairs when s is sufficiently large (towards $+\infty$ in extreme cases), and (perfect) complementary pairs when s is sufficiently small (towards 0 in extreme cases).

4.1.2 Exponential Function

Similarly, we exam the exponential function form (Tab. 1) for different values of b. When $b = 0$, the resulted MRS is constant $\frac{a}{1-a}$; when b goes to ∞, the resulted MRS goes to infinity when $q_j > q_k$ and drops to zero when $q_j < q_k$. These suggest that the exponential functional form can capture complements and substitutes.

Solving the differential equation Eq. 3 yields:

$$ae^{-bq_k} + (1-a)e^{-bq_j} = const. \tag{6}$$

The corresponding utility function is,

$$U(q_j, q_k) = z\left(ae^{-bq_k} + (1-a)e^{-bq_j}\right) \tag{7}$$

4.2 Multi-product Utility Modeling

In practice, it is very common that there are more than two products in a single transaction/order and it's desirable for us to represent the utility of arbitrary number of products. Let Ω_{it} be the set of products purchased by user i at time t. We consider the utility of Ω_{it} as the sum of the utility of all product pairs within Ω_{it}, namely,

$$\begin{aligned} U(\Omega_{it}) &= \tfrac{1}{|\Omega_{it}|-1} \sum_{j,k\in\Omega_{it}, j\neq k} U(q_j, q_k) \\ &= \begin{cases} \tfrac{1}{|\Omega_{it}|-1} \sum\limits_{j,k\in\Omega_{it}, j\neq k} \left(a_{ij}q_j^{1-b_{jk}} + (1-a_{jk})q_k^{1-b_{jk}}\right) & \text{for Eq.(5)} \\ \tfrac{1}{|\Omega_{it}|-1} \sum\limits_{j,k\in\Omega_{it}, j\neq k} \left(a_{ij}e^{-b_{jk}q_j} + (1-a_{jk})e^{-b_{jk}q_k}\right) & \text{for Eq.(7)} \end{cases} \end{aligned} \tag{8}$$

where a_{ij} and b_{ij} are product pair specific parameters, $|\Omega_{it}|$ is the number of products in set Ω_{it}, and $U(q_j, q_k)$ is the utility of two products described in the previous subsection.

4.3 CF-based Re-Parameterization

As seen from Eq. (8), there are two unknown parameters a_{jk}, b_{jk} for product j and k. Inspired by CF, we propose to model the parameters as below,

$$a_{jk} = \sigma\left(\alpha + \beta_j + \beta_k + \vec{x}_j^T\vec{x}_k\right) \tag{9}$$

$$b_{jk} = \exp\left(\mu + \gamma_j + \gamma_k + \vec{p}_j^T\vec{p}_k\right) \tag{10}$$

$$\vec{x}_j, \vec{p}_j \in \mathbb{R}^d, \beta_j, \gamma_j, \alpha, \mu \in \mathbb{R}$$

where $\sigma(\cdot)$ is Sigmoid function that ensures $0 < a_{jk} < 1$ and exponential function ensures $b_{ij} > 0$. Under CF representation, the parameters now are $\Theta = \{\vec{x}_j, \vec{p}_j, \beta_j, \gamma_j, \alpha, \mu\}$.

4.4 Discrete Choice Modeling

In economics, *discrete choice models* characterize and predict consumer's choices between two or more alternatives, such as buying Coke or Pepsi, or choosing between different hotels for traveling. In this paper, at each time point t, consumer chooses product set Ω_{it} over multiple alternatives. Let $g(\Omega_{it})$ represents other alternative candidate products set for a chosen Ω_{it}. Let $\Pi_{it} = \{\Omega_{it}, g(\Omega_{it})\}$ to represent all product sets and its k-th element is Π_{it}^k. Researchers in economics have developed *random utility models* (RUMs) for the discrete choice problem[17]. RUMs attach each alternative utility with a random value:

$$U_i(\Pi_{it}^k)' = U_i(\Pi_{it}^k) + \epsilon_k \tag{11}$$

where ϵ_k is a random variable that follows a certain probability distribution. The probability that a consumer chooses Π_{it}^1 (i.e. Ω_{it}) over other alternatives is:

$$P\left(U_i(\Pi_{it}^1)' > U_i(\Pi_{it}^k)'\right) = P\left(\epsilon_k - \epsilon_1 < U_i(\Pi_{it}^1) - U_i(\Pi_{it}^k)\right) \tag{12}$$

where $k = 2, \ldots, |\Pi_{it}|$. If ϵ_1 and ϵ_k follow *iid* extreme value distribution, it can be shown that the probability of choosing Π_{it}^1 is the following multinomial logistic model (MNL):

$$P(y_{it} = 1) = \frac{\exp\left(U_i(\Pi_{it}^1)\right)}{\sum_{k=1}^{|\Pi_{it}|} \exp\left(U_i(\Pi_{it}^k)\right)} \tag{13}$$

Alternatively, if ϵ_k follows a Gaussian distribution, $P(y_{it})$ becomes a Probit models. In the rest of this paper, we will use multinomial logistic regression.

At each time point for a given users, the system usually observes a chosen product set (e.g. an order with multiple products, a wishlist) Ω_{it}, while negative product sets are not observed. We can construct alternative sets $g(\Omega_{it})$ as negative training data using sampling strategies.

4.5 Budget Constraint

The theory of consumer choice in microeconomics[9] is concerned about how consumers maximize their utility of their consumption subject to their budget constraint. The utility of consumption is determined by consumers preference and their corresponding utility mode as explained in Section 4.2. In economics, the consumer choice problem is formalized as the following constrained optimization problem,

$$\begin{aligned} &\underset{\{q_1, q_2, \ldots, q_N\}}{\operatorname{argmax}} \quad U_{it}(q_1, q_2, \ldots, q_N) \\ &\qquad s.t. \sum_{j=1}^{N} p_j \times q_j \leq W_{it} \end{aligned} \tag{14}$$

where p_j is the price of product j, q_j is the consumed quantity of product j, and W_{it} is the consumer's budget. The solution of Eq. (14) can be obtained by standard constraint optimization methods if the quantity variables q_j are real numbers. However, q_j are discrete numbers in most of the cases, this turns the the above optimization problem into an integer programming problem, which is NP hard. Due to the exponential computational complexity, it is not feasible to consider all possible product combinations for the objective function in Eq. (14). When training the utility model, we only generative a sample of candidate sets Π_{it} for each observed chosen product set Ω_{it}.

4.6 Model Parameter Learning

Given the observed transactions/orders and the consumer discrete choice modeling framework, the model parameters Θ can be optimized by maximizing the following log-likelihood of training data:

$$\begin{aligned} &\underset{\Theta}{\operatorname{argmax}}\, nll(D; \Theta) \\ &= \sum_{i,t:I_{i,t}=1} \log\left(P(y_{it} = 1)\right) + \eta ||\Theta||^2 \end{aligned} \tag{15}$$

where D is the training dataset. $I_{it} = 1$ if user i places an order at time t. $P(y_{it})$ is the multinomial logistic regression model described in Section 13. $\eta \in \mathbb{R}^+$ is regularization coefficient which is determined using cross validation.

There is no close form solution, the optimal model parameters can be found using gradient based methods such as stochastic gradient descent.

5. MULTI-PRODUCT RECOMMENDATION

The objective of our recommendation algorithm is to recommend a set of products that gives the maximum utility without violating the budget constraint, as defined in Eq. (14). As we will see later, the purchase quantity of each product for a given user can be predicted. Eq. (14) can be reformulated as,

$$\underset{\Omega_{it}}{\operatorname{argmax}}\, U(\{q_j | j \in \Omega_{it}\}) \tag{16}$$

$$s.t. \sum_{j \in \Omega_{it}} p_j \times q_j \leq W_{it} \tag{17}$$

where Ω_{it} is a subset of all products. In practice, it is reasonable to limit $|\Omega_{it}|$ based on the typical size of an order. Due to the large search space of candidate products, it is not computationally feasible to evaluate all sets exhaustively. Instead, we resort to some heuristic approach that gives an approximate solution. For example, one idea is to grow Ω_{it} incrementally by adding a product that gives the maximum incremental utility.

6. EXPERIMENT

We studied the proposed framework based on two real world E-commerce datasets. The experimental design and results are reported in this Section.

6.1 Dataset Description

The following two real-world datasets are used in our experiments:

Table 3: Basic statistics of the two datasets

Dataset	#Transactions	#Products	Average Size	Train/Test
***.com	86k	370k	~ 8	80%/20%
Amazon	7.8k	18k	~ 12	

****.com Data* : Each record in the dataset is a purhcasing transaction with consumer id, product(s) price, product(s) quantity and the purchasing time. We treat each transaction as a positive training data point for Equation 13. The key data statistics is summarized in Table 3. As we are focusing on multiple products, we processed the dataset by removing transactions with less than two products.

Amazon Baby Registry Data Amazon's Baby Registry[1] allows consumer to add and manage products for babies. Each registry is like a wishlist which contains a list of products the list owner wants to purchase. As the lists are publicly available, we crawled the lists and their products to generate this data set. Each product comes with title, price, brand, category information. Some of the key statistics of the dataset is summarized in Table 3. We treat each wish list a positive training point for Equation 13.

Each dataset involved can be viewed as a collection of transactions. Each transaction holds a set of products consumer purchased or wanted at certain time. The transactions are randomly split into two subsets - 80% of them are used for model training and the rest 20% is for performance evaluation. For each transaction, a small portion (20%) of the products are randomly masked and they are predicted by recommendation algorithm based on other observed products in the same transaction.

[1] https://www.amazon.com/babyregistry

Table 2: Evaluation results for Top-K recommendation performance on Precision, Recall, and F_1-measure.

Dataset	Amazon Baby Registry Transactions								
@K	1			5			10		
Method	CF	BPRMF	MPUM	CF	BPRMF	MPUM	CF	BPRMF	MPUM
Precision (%)	0.092	0.117	**0.275**	0.437	0.473	**0.609**	0.262	0.513	**0.669**
Recall (%)	0.074	**0.112**	0.103	0.761	0.866	**1.279**	1.178	1.150	**2.844**
F_1-measure	0.082	0.114	**0.150**	0.555	0.612	**0.825**	0.429	0.710	**1.083**

Dataset	***.com Transactions								
@K	1			5			10		
Method	CF	BPRMF	MPUM	CF	BPRMF	MPUM	CF	BPRMF	MPUM
Precision (%)	0.022	0.076	**0.470**	0.012	0.038	**0.286**	0.012	0.026	**0.160**
Recall (%)	0.017	0.003	**0.465**	0.035	0.013	**1.390**	0.073	0.185	**1.531**
F_1-measure	0.019	0.006	**0.467**	0.018	0.019	**0.474**	0.021	0.046	**0.290**

For the first 80% training transactions, we generate negative training data (i.e. product sets not chosen by a user) for each positive set, as they are required in Eq.13. For computational efficiency, we only generate negative product sets closer to the target positive chosen set. Given a chosen product set (an order or a wishlist), we assume the budget is the total cost of the products in the chosen set. We keep the products unchanged and enumerate all quantity combinations that are subject to the same budget) constraint. Each quantity combination acts as a purchasing alternative. For computational efficiency, we further limit the size of Π_{it} to be 10 by randomly sampling from $g(\Omega_{it})$.

6.2 Evaluation Metric

Precision and recall at top-K are used for evaluation, as they are the most widely used ranking evaluation metrics in existing literature. Let Γ_i be the masked items in the i-th testing transaction and Γ_i' is a list of recommended items by the recommendation algorithm under consideration. The metrics are defined as follows:

$$\begin{aligned} \text{Precision@K} &= \frac{1}{N}\sum_{i=1}^{T}\frac{|\Gamma_i' \cap \Gamma_i|}{K} \\ \text{Recall@K} &= \frac{1}{N}\sum_{i=1}^{T}\frac{|\Gamma_i' \cap \Gamma_i|}{|\Gamma_i|} \\ F_1\text{-measure@K} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned} \tag{18}$$

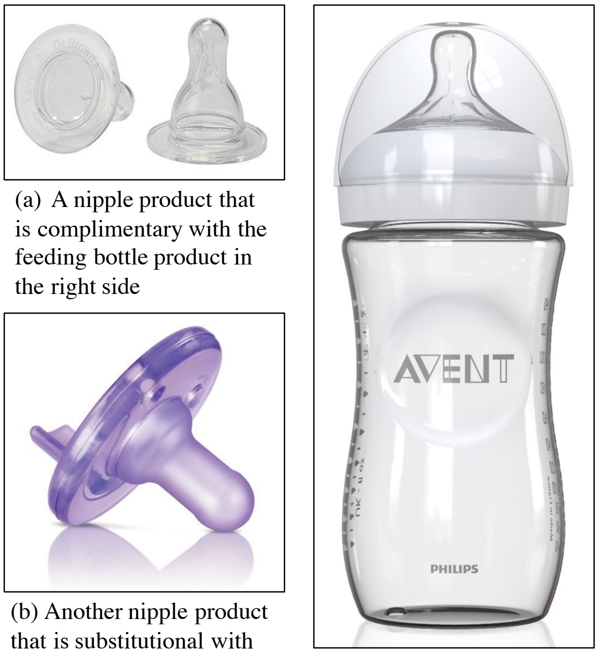
where K is the length of the recommendation list.

6.3 Experimental Results

We investigated the performance of our MPUM framework for the task of product recommendation for a transaction. For performance comparison, we considered CF based algorithm described in Section 4.3 and Bayesian Personalized Ranking (BPR) [20] as the baseline algorithms. Both CF and BPR recommend by predicting the purchasing quantities directly. $|\Pi_{it}|$ in Eq. (13) is set to 10 and SGD learning rate is set to 0.01. For fair comparison, shared parameters of different models are set to be the same: the latent factor size is set to 10 and the regularization coefficient η are set to be 0.01.

The evaluation results on Amazon and ***.com datasets are reported in Table 2, and the largest value on each dataset and for each evaluation measure is significant at 0.01 level.

It can be seen from the results that our proposed MPUM algorithm outperforms the baseline algorithms in nearly all the cases, and in particular, the performance advantage is more pronounced on ***.com dataset. A possible reason is that ***.com dataset has much lower density (0.00205%) than Amazon dataset (0.0655%). Compared to baseline algorithms, our method is less sensitive by low density. This is because that the CF and BPR approaches introduce latent vectors for users (i.e., transactions in our problem) and products, and then learn the vectors through user-product interaction pairs; while our MPUM algorithm only concerns product-product relationships and models the transactions indirectly through its products without the need to considering the vastly sparse user-product pairs, as a result, our MPUM requires much less model parameters than the baseline algorithms.



(a) A nipple product that is complimentary with the feeding bottle product in the right side

(b) Another nipple product that is substitutional with other nipple products

(c) A feeding bottle product

Figure 2: Real-world examples of complementary and substitutional products from Amazon Baby Registry dataset.

6.4 Further Analysis: Empirical Study of Economic Intuition

We did some further analysis to investigate the economic intuition of our approach in terms of the learned utility functions. In our analysis, we focus on the CES utility function in Eq.(5), because by examining the Elasticity of Substitution (ES) for real-world products learned by our model, we hope to find intuitive explanations for our principled economic-driven approach in practical applications.

We look at some the real-world product pairs with the lowest (complementary) and highest (substitutional) ES values in our model. As shown in Figure 2, we find that the product pair with the lowest ES is a nipple together with a feeding bottle (Figure 2(a) and 2(c)), which are clearly complementary products. The pair with the highest ES are two different brands of nipple products (Figure 2(a) and 2(b)), which are substitutes because users usually only needs to purchase either one of the two.

We also compute the average elasticity of substitution for each product in the Amazon Baby Registry dataset by averaging its estimated ES with all other products. We find that the popularity of a product in the dataset is highly negatively correlated with the corresponding ES. This means popular products have relatively smaller ES values, which suggests popular items tend to be more complementary with other products.

More specifically, Figure 3 shows the logarithm of popularity of a product (y axis) against the average ES of the product (x axis). The correlation between log(*popularity*) and ES values is -0.916 for these products. Because we care more about the product ranking lists for recommendation rather than the absolute ES values in practice, we further rank the products according to ES and investigate the relation between log(*popularity*) and the rankings (Figure 3, right). The correlation is -0.931. Further analysis show that the products with small average ES values in Figure 3 are mostly baby care necessities (e.g., pacifier, plug, and teether) that are generally complementary with many products, which makes them generally popular in most of the transactions.

These findings are encouraging and suggest that our proposed utility maximization approach conforms with human intuitions. It makes it possible to discover product substitutional/complementary relationships from real-world transaction data automatically, based on combining machine learning techniques with principled economic theories.

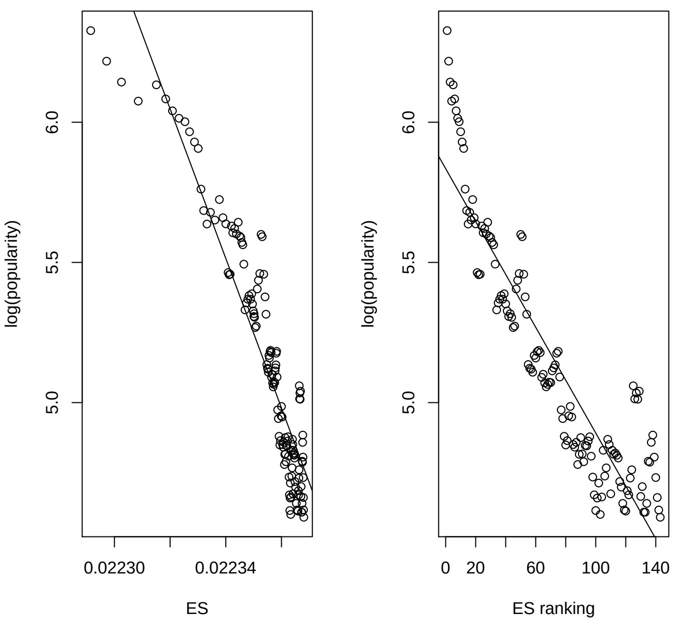


Figure 3: Scattered relations of the product popularity v.s. the average Elasticity of Substitution (ES) of the corresponding product as well as the the ranking of ES values.

7. CONCLUSIONS AND FUTURE WORK

Utility is commonly used by economists to characterize consumer preference over alternatives and it serves as cornerstone for consumer choice theory[9]. Motivated by existing research in economics, we introduced a general utility based framework for multiple products recommendation. Start with Marginal Rate of Substitution defined over products indifference curve, we derived several candidate utility functional forms that can model both substitutes and complements. The model parameters are learned based on existing consumer data. Recommendations of multiple products are generated by maximizing the learned utility model. Experimental results on both Amazon and ***.com e-commerce data sets demonstrated the effectiveness of the proposed approach for recommendation. Further analysis also shows complements and substitutes found by the model look reasonable.

Modeling the relationships between products is a fundamental problem for various recommendation tasks, such as package recommendation, next basket recommendation and top K products recommendation. Although our experiments are about top K products recommendation, the proposed framework can be applied to other usage scenario in the future. We expect the proposed framework complements some very different existing methods that implicitly captures products relationships such as list-wise CF or list-wise learning to rank, and it would be interesting to compare them in the future. This is a first toward multi-products utility modeling and there are much room to further improve the techniques. For example, the functional form of MRS could be adjusted to capture other products relationships besides complements and substitutes. We can also introduce products features and users features into this framework. We used a greedy method to generate top K products, because maximize the utility function learned is an integer linear programming problem and is NP-hard. Other heuristic methods that have been applied to 0-1 integer programming problems can tried in the future.

8. REFERENCES

[1] Y. K. Adomavicius, G. Toward more diverse recommendations: Item re-ranking methods for recommender systems. In *Proc. of the 19th Workshop on Information Technologies and Systems*, pages 657–666. ACM, 2009.

[2] E. Brynjolfsson, Y. J. Hu, and M. D. Smith. Research commentary—long tails vs. superstars: The effect of information technology on product variety and sales concentration patterns. *Info. Sys. Research*, 21(4):736–747, Dec. 2010.

[3] R. E. Bucklin and V. Srinivasan. Determining interbrand substitutability through survey measurement of consumer preference structures. *Journal of Marketing Research*, pages 58–71, 1991.

[4] J. Canny. Collaborative filtering with privacy via factor analysis. In *SIGIR*, pages 238–245, 2002.

[5] P. Castells, J. Wang, R. Lara, and D. Zhang. Workshop on novelty and diversity in recommender systems. 2011.

[6] K. Christidis, D. Apostolou, and G. Mentzas. Exploring customer preferences with probabilistic topic models. In *ECML*, 2010.

[7] J. S. Coleman and T. J. Fararo. Rational Choice Theory. *Nueva York: Sage*, 1992.

[8] M. D. Ekstrand, J. T. Riedl, and J. A. Konstan. Collaborative Filtering Recommender Systems. *Foundations and Trends in Human-Computer Interaction*, 4(2):81–173, 2011.

[9] J. Gans, S. King, and N. G. Mankiw. Principles of Microeconomics. *Cengage Learning*, 2011.

[10] A. Gatzioura and M. Sanchez-Marre. A case-based recommendation approach for market basket data. *Intelligent Systems, IEEE*, 30(1):20–27, 2015.

[11] N. Hurley and M. Zhang. Novelty and diversity in top-n recommendation–analysis and evaluation. *Transactions on Internet Technology (TOIT)*, 10(4):14, 2011.

[12] Y. Koren, R. Bell, and C. Volinsky. Matrix Factorization Techniques for Recommender Systems. *Computer*, 2009.

[13] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[14] Y. J. Lim and Y. W. Teh. Variational Bayesian approach to movie rating prediction. In *Proceedings of KDD Cup and Workshop*, 2007.

[15] P. Lops, M. De Gemmis, and G. Semeraro. Content-based recommender systems: State of the art and trends. In *Recommender systems handbook*. Springer, 2011.

[16] J. J. McAuley, R. Pandey, and J. Leskovec. Inferring networks of substitutable and complementary products. In *KDD*, 2015.

[17] D. McFadden. Econometric models for probabilistic choice among products. *Journal of Business*, pages S13–S29, 1980.

[18] A. Paterek. Improving regularized singular value decomposition for collaborative filtering. In *Proceedings of KDD Cup and Workshop*, 2007.

[19] P. Pu, L. Chen, and R. Hu. A user-centric evaluation framework for recommender systems. In *RecSys*, pages 157–164. ACM, 2011.

[20] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pages 452–461. AUAI Press, 2009.

[21] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 811–820. ACM, 2010.

[22] J. D. M. Rennie and N. Srebro. Fast maximum margin matrix factorization for collaborative prediction. *ICML*, pages 713–719, 2005.

[23] M. C. Rios, C. R. McConnell, and S. L. Brue. *Economics: Principles, problems, and policies.* McGraw-Hill, 2013.

[24] G. J. Russell and R. N. Bolton. Implications of market structure for elasticity structure. *Journal of Marketing Research*, pages 229–241, 1988.

[25] G. J. Russell and A. Petersen. Analysis of cross category dependence in market basket selection. *Journal of Retailing*, 76(3):367–392, 2000.

[26] R. Salakhutdinov and A. Mnih. Bayesian probabilistic matrix factorization using markov chain monte carlo. *ICML*, pages 880–887, 2008.

[27] B. Sarwar, G. Karypis, et al. Item-based collaborative filtering recommendation algorithms. *WWW*, 2001.

[28] J. B. Schafer, J. Konstan, and J. Riedl. Recommender systems in e-commerce. In *EC*, pages 158–166, 1999.

[29] Y. Shi, X. Zhao, J. Wang, M. Larson, and A. Hanjalic. Adaptive diversification of recommendation results via latent factor portfolio. In *SIGIR*, SIGIR '12, pages 175–184, New York, NY, USA, 2012. ACM.

[30] A. D. Shocker, B. L. Bayus, and N. Kim. Product complements and substitutes in the real world: The relevance of ?other products? *Journal of Marketing*, 68(1):28–40, 2004.

[31] G. Takacs, I. Pilaszy, B. Nemeth, and D. Tikk. Investigation of Various Matrix Factorization Methods for Large Recommender Systems. *Proc. ICDM*, 2008.

[32] J. Wang and Y. Zhang. Utilizing marginal net utility for recommendation in e-commerce. In *SIGIR*, pages 1003–1012. ACM, 2011.

[33] P. Wang, J. Guo, and Y. Lan. Modeling retail transaction data for personalized shopping recommendation. In *CIKM*, pages 1979–1982, 2014.

[34] P. Wang, J. Guo, Y. Lan, J. Xu, S. Wan, and X. Cheng. Learning hierarchical representation model for nextbasket recommendation. In *SIGIR*, 2015.

[35] Y. Zhang, Q. Zhao, Y. Zhang, D. Friedman, and M. Zhang. Web intelligence for social good: Recommendation based on total surplus maximization. *WWW*, 2015.

[36] Q. Zhao, Y. Zhang, D. Friedman, and F. Tan. E-commerce Recommendation with Personalized Promotion. *RecSys*, pages 219–226, 2015.

[37] T. Zhou, Z. Kuscsik, J.-G. Liu, M. Medo, J. R. Wakeling, and Y.-C. Zhang. Solving the apparent diversity-accuracy dilemma of recommender systems. *National Academy of Sciences*, 107(10), 2010.

Discussion Papers of the Research Area Markets and Choice 2016

Research Unit: **Market Behavior**

David Danz, Steffen Huck, Philippe Jehiel Public statistics and private experience: Varying feedback information in a take-or-pass game	SP II 2016-201
Jana Friedrichsen Signals sell: Designing a product line when consumers have social image concerns	SPII 2016-202
Uri Gneezy, Silvia Saccardo, Roel van Veldhuizen Bribery: Greed versus reciprocity	SPII 2016-203
Inácio Bó, C.-Philipp Heller Strategic schools under the Boston mechanism revisited	SPII 2016-204

Research Unit: **Economics of Change**

Armin Falk, Nora Szech Pleasures of skill and moral conduct	SP II 2016-301
Thomas Deckers, Armin Falk, Fabian Kosse, Nora Szech Homo moralis: Personal characteristics, institutions, and moral decision-making	SP II 2016-302
Jenny Simon, Justian Mattias Valasek The political economy of multilateral aid funds	SP II 2016-303
Vittorio Bassi, Steffen Huck, Imran Rasul A note on charitable giving by corporates and aristocrats: Evidence from a field experiment	SP II 2016-304
Ludwig Ensthaler, Steffen Huck, Johannes Leutgeb Games played through agents in the laboratory – A test of Prat & Rustichini's model	SP II 2016-305
Maja Adena, Steffen Huck Online fundraising, self-deception, and the long-term impact of ask avoidance	SP II 2016-306
Aniol Llorente-Saguer, Roman M. Sheremeta, Nora Szech Designing contests between heterogeneous contestants: An experimental study of tie-breaks and bid-caps in all-pay auctions	SP II 2016-307

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice

Research Professorship: **Market Design: Theory and Pragmatics**

Kevin McLaughlin, Daniel Friedman SP II 2016-501
Online ad auctions: An experiment

Yongfeng Zhang, Qi Zhao, Yi Zhang, Daniel Friedman, Min Zhang, Yiqun Liu, Shaoping Ma SP II 2016-502
Economic recommendation with surplus maximization

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice