POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

An Online Dynamic Pricing Algorithm for Complementary Products

Tesi di Laurea Magistrale in
Computer Science and Engineering

Author: **Andrea Cerasani**

Student ID: 103375
Advisor: Prof. Marcello Restelli
Co-advisors: Marco Mussi, Alessandro Lavelli
Academic Year: 2022-23

Abstract

The traditional pricing paradigms, once reliant on static models and rule-based strategies, rapidly give way to dynamic data-driven approaches powered by machine learning algorithms. Dynamic pricing algorithms usually face the problem of finding the optimal prices of a product independently from the others. However, this choice may lead to suboptimal solutions, as we miss the chance to exploit product interactions. In this thesis, we present `CPP` (`Complementary Product Pricing`), an online learning algorithm for optimizing the pricing strategies of a set of products, considering the substitutable and complementary relations between them. The algorithm makes use of transaction data to learn the interaction between the different items and then optimize the pricing strategies through efficient multi-armed bandit solutions. We validate our solution in a simulated environment that mimics the one of a real e-commerce website, and we demonstrate that `CPP` improves the profit w.r.t. to an algorithm that ignores such interaction, also in the short-term, up to 30%.

Keywords: complementary products, dynamic pricing, multi-armed bandit, online learning

Abstract in lingua italiana

I paradigmi tradizionali di pricing, un tempo basati su modelli statici e strategie rule-based, stanno lasciando rapidamente il posto ad approcci dinamici basati sui dati e alimentati da algoritmi di apprendimento automatico. Gli algoritmi di pricing dinamico di solito affrontano il problema di trovare i prezzi ottimali di un prodotto indipendentemente dagli altri. Tuttavia, questa scelta può portare a soluzioni subottimali, poiché si perde la possibilità di sfruttare le interazioni tra i prodotti. In questa tesi presentiamo `CPP` (`Complementary Product Pricing`), un algoritmo di apprendimento online per ottimizzare le strategie di prezzo di un insieme di prodotti, considerando le relazioni di sostituibilità e complementarietà tra di essi. L'algoritmo utilizza i dati delle transazioni per apprendere l'interazione tra i diversi articoli e quindi ottimizzare le strategie di prezzo attraverso soluzioni efficienti con multi-armed bandit. Validiamo la nostra soluzione in un ambiente simulato che imita quello di un sito web di un e-commerce reale e dimostriamo che `CPP` migliora il profitto rispetto a un algoritmo che ignora tale interazione, anche nel breve termine, fino al 30%.

Parole chiave: prodotti complementari, pricing dinamico, multi-armed bandit, apprendimento online

Contents

1 | Introduction

Determining the optimal price of a product is a crucial task for retailers, playing a significant role in shaping consumer behaviour, influencing purchasing decisions, and ultimately impacting the overall success of a retail business. This task is directly addressed by *dynamic pricing*, the study of determining optimal selling prices of products or services, in a setting where prices can easily and frequently be adjusted (Den Boer, 2015). Among the key factors that determine the demand for a product and its optimal selling price, the relations each product has with others sold by the same retailer is certainly of great importance. Studies demonstrate the inherent value in pricing by adopting bundling and joint pricing strategies (Yan and Bandyopadhyay, 2011; Venkatesh and Kamakura, 2003). Dynamic pricing algorithms usually face the problem of finding the optimal prices of a product independently from the others, which may lead to suboptimal solutions, as we miss the chance to exploit product interactions. Two main types of relationships can be identified between products: substitutability and complementarity. Substitutable products refer to goods that can be used in place of each other to satisfy a particular need or want. As a consequence, these products are rarely bought together, i.e. two different brands of mattresses or two different models of computer keyboards. Complementary products, on the other hand, are products that "go together" (Nicholson and Snyder, 2017) and complement each other, such as a printer and ink cartridges or a toothbrush and toothpaste. They enhance the value and utility of each other when consumed or utilized simultaneously and thus are usually bought together. This characteristic of complementary products is greatly exploited in recommender systems (McAuley et al., 2015) to suggest additional products for the user to buy given the interest shown in their complementary products and in pricing (Mulhern and Leone, 1991; Ghoniem et al., 2016; Feng et al., 2018), to take into account product relations when selecting the optimal selling price.

1.1. Goal and Challenges

Dynamic pricing algorithms usually face the problem of finding the optimal prices of a product independently from the others and this can lead to suboptimal solutions as we miss the chance to exploit product interactions. The goal of our thesis is to fill this gap by proposing an online learning algorithm for optimizing the pricing strategy of a set of products, considering both positive and negative interactions between them. The main challenges reside in the characteristics of the complementary relationship and in the complexity of the problem. A complementarity relationship exhibits characteristics such as asymmetry and non-transitivity (Kocas et al., 2018; Yu et al., 2019; Xu et al., 2020) which distinguish complementarity from a simple equality relation like substitutability. Moreover, two products can be complementary while belonging to semantically different areas, so complementarity is inherently harder to capture and cannot be identified simply by similarity. When determining the price of a product, we are dealing with an unknown environment and the selling volumes that a product will produce at a given price are not known in advance. Thus, when dynamically pricing a product, we need to balance the information we acquired about the environment and collecting new information. This is known as the exploration-exploitation dilemma and is to be addressed when dealing with such a task. Another crucial challenge to tackle is the computational complexity of the problem. Indeed, given a catalogue of products, each product can be related to all of the others in the catalogue with a combinatorial explosion that is infeasible to manage both from a computational and data scarcity point of view.

1.2. Original Contribution

In this thesis, we address these challenges, focusing on pricing non-luxury products in retail e-commerce with unlimited inventory. We propose an online learning algorithm to identify complementary products and optimize their pricing strategies by taking into account the interactions between them. The algorithm makes use of transactional data to identify the complementarity relationships between products. The computational complexity of the problem is tackled by considering only the most significant complementary relations among products. A Multi-Armed Bandit approach is used to learn the product demands, addressing the exploration-exploitation tradeoff with the use of Thompson Sampling as a heuristic to drive exploration in the choice of product prices. We evaluate our solution in a simulated environment that mimics the one of a real e-commerce website, and we demonstrate an improvement of up to 30% in profit with respect to an algorithm that ignores product relations and assumes the products to be independent.

1.3. Work Structure

Our work is organized as follows: in Chapter 2 we provide an overview of the topics that are the backbone of our proposed solution. We cover Economics Groundings, Hypothesis Tests, Linear Models used in regression with a focus on Bayesian Linear Regression and Multi-Armed Bandits. In Chapter 3 we formally define the problem we want to address in the thesis and the objective. In Chapter 4 we discussed our proposed solution of the problem, describing our algorithm `CPP`. In Chapter 5, we explore works related to the problem we face in our thesis, comparing them to `CPP`. In Chapter 6 we validate our algorithm in a synthetic environment and compare it to an algorithm that independently price products. Finally, in Chapter 7 we draw the conclusions on our thesis, suggesting future extension of our work in new settings.

2 | Background

In this chapter, we aim to provide the groundings on the topics we will address in our thesis. We first introduce in Section 2.1 the groundings in microeconomics needed in our thesis, in particular the concepts of demand, complementary and substitutable products and cross-price elasticity. In Section 2.2 we cover the statistical tools we make use of in our thesis, focusing in particular on the Binomial Hypothesis Test. We proceed in Section 2.3 to discuss the mathematical formulation of Linear Basis Function Models and their extension to a Bayesian framework with Bayesian Linear Regression. In Section 2.4 we cover the Multi-Armed Bandit problem, focusing on the Bayesian approaches used to solve the exploration-exploitation dilemma, such as Thompson Sampling.

2.1. Economics Groundings

2.1.1. Demand Function and Price Elasticity

A *demand function*, in the context of microeconomics, describes the relationship between the price of a good or service and the quantity of that good or service that consumers are willing and able to purchase at different price levels. As shown in figure 2.1 we depict the demand function with the price on the x-axis and the quantity on the y-axis, in order to better convey the fact that price is considered an independent variable and quantity depends on price. Closely related to the demand function is the concept of *price elasticity of demand*, which quantifies the responsiveness of the quantity demanded to changes in the price of a good or service. For most goods, following the law of demand, the demanded quantity decreases as the price increases (Nicholson and Snyder, 2017) and accordingly, the demand function is decreasing and the price elasticity is negative. Exceptions to this case are particular categories of goods such as Veblen goods (Kemp, 1998) and Giffen goods (Dougan, 1982).

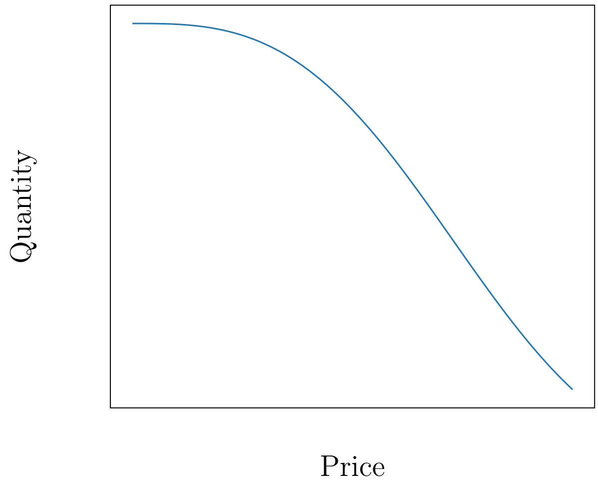


Figure 2.1: Example of a demand function

2.1.2. Complementary and substitutable products

The economic literature distinguishes between two types of product interdependencies: *complementary products* and *substitutable products*(Nicholson and Snyder, 2017).

Complementary products "go together" (Nicholson and Snyder, 2017) and complement each other. These kinds of products are usually bought together, such as a printer and an ink cartridge or a toothbrush and toothpaste.

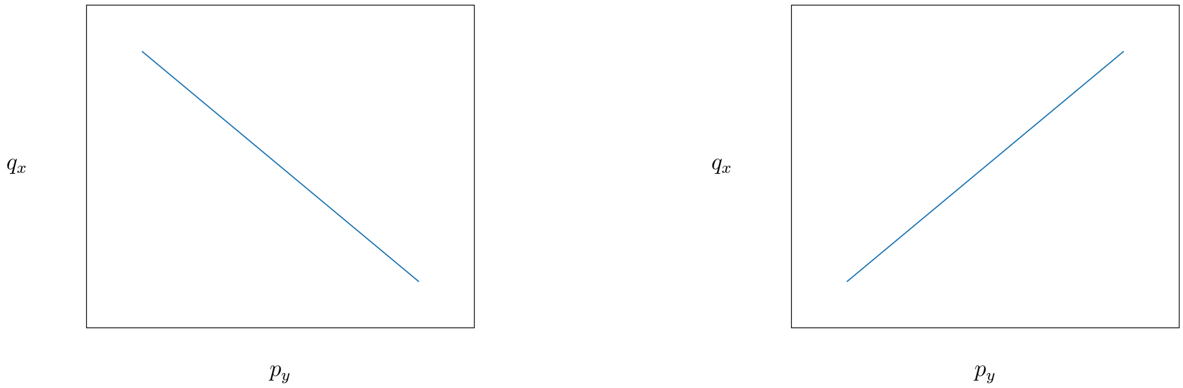
On the other hand, substitutable products are products that serve the same purpose and are rarely bought together, instead, they usually cannibalize each other, for example, two different brands of a mattress are rarely bought together and the purchase of one brand usually corresponds to one less purchase of the other brand.

2.1.3. Cross-Price Elasticity

Complementary and substitutable products have different effects on the demand function for related products. The concept of *cross-price elasticity* is useful to describe these effects: it measures the proportionate change in the quantity q of a good demanded in response to a proportionate change in the price of some other good (Nicholson and Snyder, 2017).

More formally, given two products x and y, their respective prices p_x and p_y and the demand function of product x $q_x(p_x, p_y)$, the cross-price elasticity of demand for product x with respect to the price of product y is defined as:

$$e_{xy} = \frac{\partial}{\partial p_y} q_x(p_x, p_y) * \frac{p_y}{q_x}$$



(a) Complementary products

(b) Substitutable products

Figure 2.2: Cross price elasticities between products

For *complementary products*, as we can see in figure 2.2a the cross-price elasticity is negative: this is intuitive since a lower price of product y leads to a lower quantity of product y and, since the demand of product y drives the demand of the complementary product x, we can expect also the quantity of product x to decrease.

On the contrary, as shown in figure 2.2b, the effect for *substitutable products* is the opposite: the cross-price elasticity is positive since the lower demand of product y leads to an increase in the demand of a competing product x.

2.2. Hypothesis Test

A statistical hypothesis is a statement about a set of parameters of a population distribution. A primary problem is to develop a procedure for determining whether or not the values of a random sample from this population are consistent with the hypothesis (Ross, 2021).

2.2.1. Significance levels

Given a population having distribution F_θ, where θ is unknown, we want to test a specific hypothesis about θ. We denote this hypothesis by H_0 and call it the *null hypothesis*.

In order to test a specific null hypothesis H_0, a population sample of size n $X_1, \ldots, X_n$ is observed. The decision on whether or not to accept H_0 is based on these n values. We can specify a test for H_0 by defining a region C in n-dimensional space. If the random sample $X_1, \ldots, X_n$ lies in C the hypothesis is rejected, accepted otherwise. The region C is called the *critical region*. We encounter two types of errors in testing, given a null hypothesis H_0: *type I error* is said to result if the test rejects H_0 when it is correct while *type II error* results if the test accepts H_0 when it is false.

The objective of a statistical test is not to explicitly determine if the null hypothesis is true or not but rather its consistency with the resultant data. Thus, we want to reject H_0 only with very unlikely data. To accomplish this, we specify a value α and require that whenever H_0 is true, the probability of rejecting H_0 is never greater than $alpha$. The value α is called *significance level* and we usually set it in advance and common values for it are $\alpha = \{0.1, 0.05, 0.005\}$.

We can determine whether or not to accept the null hypothesis by computing first the value of the test statistic, i.e. a function $d(X_1, \ldots, X_n)$ that maps samples to a value used to accept or reject the null hypothesis, and compute the probability that the distribution used in the test would exceed that quantity. This probability is called the p-value and gives the critical significance level in the sense that H_0 will be accepted if the significance level α is less than the p-value and rejected if it is greater than or equal.

2.2.2. Binomial Test

The binomial test is a test of the null hypothesis H_0 that the probability of success p in a binomial distribution is p_0. Considering a binomial distribution with parameters (n, p), we define the following testing problem:

$$\begin{aligned} H_0 &: p \leqslant p_0 \\ H_1 &: p > p_0 \end{aligned} \tag{2.1}$$

where H_1 is called the *alternative hypothesis* and p_0 is some specified value.

This is called a *one-sided* testing problem since we call for rejection only when the estimator of p is large, as opposed to the *two-sided* case where the alternative hypothesis is $H_1 : p \neq p_0$.

If we let X denote the number of successes in the sample of size n, we wish to reject H_0 when X is large and we have that:

$$P\{X \geqslant k\} = \sum_{i=k}^{n} P\{X = i\} = \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i}. \tag{2.2}$$

It can be proven that $P\{X \geqslant k\}$ is an increasing function of p and using this we observe that when H_0 is true (so $p \leqslant p_0$),

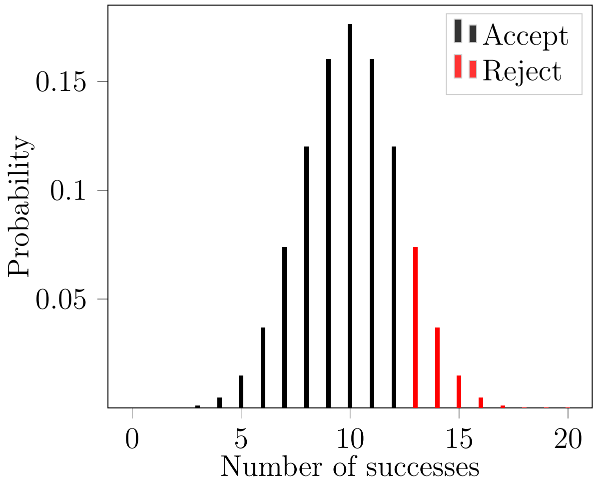


Figure 2.3: Example of acceptance and rejection regions in a binomial test

$$P\{X \geqslant k\} \leqslant \sum_{i=k}^{n} \binom{n}{i} p_0^i (1-p_0)^{n-i}. \tag{2.3}$$

Hence, a significance level α test of $H_0 : p \leqslant p_0$ versus $H_1 : p > p_0$ is to reject H_0 when

$$X \geqslant k^* \tag{2.4}$$

where k^* is the smallest value of k for which $\sum_{i=k}^{n} \binom{n}{i} p_0^i (1-p_0)^{n-i} \leqslant \alpha$, so

$$k^* = \min \left\{ k : \sum_{i=k}^{n} \binom{n}{i} p_0^i (1-p_0)^{n-i} \leqslant \alpha \right\}. \tag{2.5}$$

We can perform this test by determining the value of the test statistic, e.g. $X = x$, and computing the p-value given by

$$p\text{-value} = P\{B(n, p_0) \geqslant x\} = \sum_{i=x}^{n} \binom{n}{i} p_0^i (1-p_0)^{n-i} \tag{2.6}$$

In Figure 2.3 we show a visual example of acceptance and rejection regions in a binomial test.

2.3. Linear Models

We now discuss linear models used for regression. The goal of regression is to predict the value of one or more continuous target variables t given the value of a D-dimensional vector x of input variables (Bishop, 2006). We indicate a training dataset of N observations as $\{\mathbf{x}_n\}$, where $n = \{1, \ldots, N\}$ and $\mathbf{x} = (x_1, \ldots, x_D) \in \mathbb{R}^D$ and the corresponding target values $\{\mathbf{t}_n\}$. Regression aims at modelling the relationship between target values and observations.

2.3.1. Linear Basis Function Models

The simplest linear model for regression is one that involves a linear combination of the input variables

$$y(\mathbf{x}, \mathbf{w}) = w_0 + w_1 x_1 + \ldots + w_D x_D \tag{2.7}$$

This is often simply known as *linear regression*, as it is a linear function of the parameters $\mathbf{w} = (w_0, \ldots, w_D)^T$ and, in this case, it is also a linear function of the input variables x_i, significantly limiting the expressiveness of the model. We extend the class of models by considering linear combinations of fixed non-linear functions of the input variables, of the form

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=1}^{M} w_j \phi_j(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}) \tag{2.8}$$

where $\mathbf{w} = (w_1, \ldots, w_M)^T$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_M)^T$.

The model has now M parameters where w_1 is the bias parameter with the dummy basis function $\phi_1(\mathbf{x}) = 1$.

2.3.2. Maximum Likelihood

We assume that the target variable t is given by a deterministic function $y(\mathbf{x}, \mathbf{w})$ with additive Gaussian noise so that:

$$t = y(\mathbf{x}, \mathbf{w}) + \epsilon \tag{2.9}$$

where ϵ is a Gaussian random variable with zero mean and precision (inverse variance) β.

This implies a distribution over the target variable:

$$p(t|\mathbf{x}, \mathbf{w}, \beta) = \mathcal{N}(t|y(\mathbf{x}, \mathbf{w}), \beta^{-1}). \tag{2.10}$$

Now we consider a dataset of inputs $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ with corresponding target values $\mathbf{t} = (t_1, \dots, t_N)$. With the assumption that the data points are drawn independently from the distribution, we can obtain the following expression for the likelihood function:

$$p(\mathbf{t}|\mathcal{X}, \mathbf{w}, \beta) = \prod_{n=1}^{N} \mathcal{N}(t_n|\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n), \beta^{-1}) \tag{2.11}$$

From now on we will omit the explicit $\mathcal{X}$ as it will always appear in the set of conditioning variables. We apply the logarithm to the likelihood function:

$$\begin{aligned} \ln p(\mathbf{t}|\mathbf{w}, \beta) &= \sum_{n=1}^{N} \ln \mathcal{N}(t_n|\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n), \beta^{-1}) \\ &= \frac{N}{2} \ln \beta - \frac{N}{2} \ln(2\pi) - \beta E_D(\mathbf{w}) \end{aligned} \tag{2.12}$$

where E_D is the sum-of-squares error function, defined as:

$$E_D(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} \{t_n - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n)\}^2 \tag{2.13}$$

We want to obtain the values $\mathbf{w}_{ML}$ of the parameters that maximize the likelihood of the target. To do so we compute the gradient of the log-likelihood:

$$\nabla \ln p(\mathbf{t}|\mathbf{w}, \beta) = \sum_{n=1}^{N} \{t_n - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n)\} \boldsymbol{\phi}(\mathbf{x}_n)^T. \tag{2.14}$$

We define the *design matrix* $\mathbf{\Phi}$, an $N \times M$ matrix whose elements are given by $\phi_{nj} = \phi_j(\mathbf{x}_n)$ so that

$$\mathbf{\Phi} = \begin{pmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \dots & \phi_M(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \dots & \phi_M(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \dots & \phi_M(\mathbf{x}_N) \end{pmatrix} \tag{2.15}$$

and setting the gradient to zero and solving for $\mathbf{w}$ we obtain

$$\mathbf{w}_{ML} = (\mathbf{\Phi}^T \mathbf{\Phi})^{-1} \mathbf{\Phi}^T \mathbf{t}. \tag{2.16}$$

Moreover, maximizing the log-likelihood function (2.12) with respect to β we obtain:

$$\frac{1}{\beta_{ML}} = \frac{1}{N}\sum_{n=1}^{N}\{t_n - \mathbf{w}_{ML}^T\boldsymbol{\phi}(\mathbf{x}_n)\}^2 \tag{2.17}$$

2.3.3. Bayesian Linear Regression

In order to move to a Bayesian treatment of linear regression, we introduce a prior probability distribution over the model parameters $\mathbf{w}$:

$$p(\mathbf{w}|\alpha) = \mathcal{N}(\mathbf{w}|\mathbf{m}_0, \mathbf{S}_0) \tag{2.18}$$

with mean $\mathbf{m}_0$ and covariance $\mathbf{S}_0$.

We then compute the posterior distribution, which, due to the choice of a Gaussian prior, is Gaussian as well:

$$p(\mathbf{w}|\mathbf{t}) = \mathcal{N}(\mathbf{w}|\mathbf{m}_N, \mathbf{S}_N) \tag{2.19}$$

where the posterior mean $\mathbf{m}_N$ and covariance $\mathbf{S}_N$ are given by:

$$\mathbf{m}_N = \mathbf{S}_N(\mathbf{S}_0^{-1}\mathbf{m}_0 + \beta\boldsymbol{\Phi}^T\mathbf{t}) \tag{2.20}$$

$$\mathbf{S}_N^{-1} = \mathbf{S}_0^{-1} + \beta\boldsymbol{\Phi}^T\boldsymbol{\Phi}. \tag{2.21}$$

Since the posterior distribution is Gaussian, its mode coincides with its mean and the maximum posterior weight vector is simply given by $\mathbf{w}_{MAP} = \mathbf{m}_N$.

Bayesian Linear Regression exhibits a sequential nature in that the posterior distribution forms the prior when a new data point is observed. This property makes the use of Bayesian Linear Regression suitable in an online learning setting.

2.3.4. Bayesian Monotonic Regression

We now focus on the setting where the relationship between input and target values is known to be monotonic. We are particularly interested in this scenario because the demand function can be assumed to be monotonically decreasing in price (see Section 2.1.1).

To impose monotonicity in a Bayesian Regression setting it is possible to resort to a set of basis functions such as *Bernstein Basis Polynomial.*

The k-th Bernstein Polynomial basis function of degree M is defined as

$$\psi_k(x, M) := \binom{M}{k} x^k (1-x)^{M-k}, \quad x \in [0, 1] \tag{2.22}$$

and the regression formulation results in the following weighted combination

$$f(x) = \sum_{k=0}^{M} \psi_k(x, M)\ \beta_k = \boldsymbol{\psi}\boldsymbol{\beta} \tag{2.23}$$

where $\boldsymbol{\beta} = (b_0, \ldots, b_M)^T$ and its elements are called *Bernstein coefficients* and $\boldsymbol{\psi} = (\psi_0, \ldots, \psi_M)$. The resulting model has $M + 1$ parameters.

The function in Equation 2.23 is monotone if $\beta_0 = 0$ and $b_k \geqslant b_{k-1} \quad \forall k \in \{1, \ldots, M\}$.

With the same procedure described in McKay Curtis and Ghosh (2011) and Wilson et al. (2020), we perform a reparametrization of the regression coefficients as $\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\beta}$, where $\mathbf{A}$ is an $(M + 1) \times (M + 1)$-matrix defined as

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{pmatrix}. \tag{2.24}$$

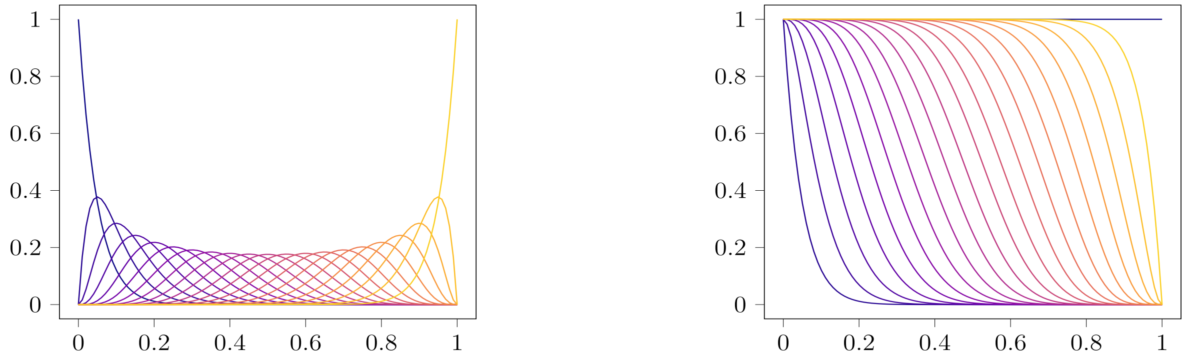
The regression function is then

$$f(x) = \boldsymbol{\psi}\mathbf{A}^{-1}\boldsymbol{\theta}. \tag{2.25}$$

With this reparametrization, f is monotone increasing when $\theta_k \geqslant 0 \quad \forall k > 0$. To obtain a monotonically decreasing function is sufficient to use as basis functions $1 - \boldsymbol{\psi}\mathbf{A}^{-1}$ instead of $\boldsymbol{\psi}\mathbf{A}^{-1}$.

In Figure 2.4 we show the Bernstein Polynomial basis functions and their transformed version.

In the context of Bayesian Linear Regression, to obtain positive $\boldsymbol{\theta}$ coefficients we resort to



(a) Bernstein Polynomial basis functions

(b) Monotonic transformation

Figure 2.4: The 21 Bernstein polynomial basis functions of degree 20

priors with non-negative support. The simplest choice is a *Lognormal* prior distribution:

$$\boldsymbol{\theta} \sim \mathcal{LN}(\boldsymbol{\theta}_0, \boldsymbol{\Sigma}_0), \tag{2.26}$$

where $\boldsymbol{\theta}_0 \in \mathbb{R}^M$ and $\boldsymbol{\Sigma}_0 \in \mathbb{R}^{M \times M}$. It is possible to estimate the posterior distribution through sampling approaches such as the *No-U-Turn Sampler* (Hoffman and Gelman, 2014) or via *Variational Inference* (Blei et al., 2017).

2.4. Multi-Armed Bandits

A bandit problem is a sequential game between an *agent* and an *environment*. The game is played over T rounds, where T is a positive natural number called the *horizon* (Lattimore and Szepesvári, 2020).

In each round $t \in \{1, \dots, T\}$, the agent chooses an action $a_t \in \mathcal{A}$ and the environment subsequently reveals a reward $r_t \in \mathbb{R}$. In literature, actions are usually called *arms* and when the number of arms is at least two, we refer to Multi-Armed Bandit (MAB).

An agent, when choosing their actions at round t, can only depend on the *history* $\mathbf{h}_{t-1} = (a_1, r_1, \dots, a_{t-1}, r_{t-1})$. A *policy* is a mapping from histories to actions and an agent adopts a policy to interact with an environment. The goal of the agent is to choose actions that lead to the largest possible cumulative reward over all the T rounds.

The fundamental challenge in bandit problems relies on the fact that the environment is unknown to the agent so the agent has to balance the acquisition of new information and the exploitation of current knowledge about the environment: this is called the *exploration-exploitation dilemma*.

2.4.1. Thompson Sampling

In order to tackle the exploration-exploitation dilemma, numerous bandit strategies have been developed over time and their mathematical properties have been extensively studied (Bubeck and Cesa-Bianchi, 2012). Thompson Sampling is a standard among the heuristics to drive exploration and its theoretical guarantees have been thoroughly researched (Agrawal and Goyal, 2012, 2013; Nuara et al., 2020). The common assumption is that the reward r is drawn from a distribution defined by parameters $\boldsymbol{\theta} \in \Theta$.

Let $P(\boldsymbol{\theta})$ be the prior distribution over the parameters and $\mathcal{D} = \{(a_i, r_i)\}_{1:t}$ the action-reward tuples up to round $t \leqslant N$. The posterior distribution of $P(\boldsymbol{\theta})$ can be inferred through Bayes' theorem:

$$P(\boldsymbol{\theta}|\mathcal{D}) \propto P(\mathcal{D}|\boldsymbol{\theta}) P(\boldsymbol{\theta}),$$

where $P(\mathcal{D}|\boldsymbol{\theta})$ is the likelihood function. With the posterior distribution of the parameters of the reward, it is possible to compute a probability distribution over the reward of each action. The strategy employed selects the action $a^* \in \mathcal{A}$ is the one with the maximum expected reward:

$$a^* = \arg\max_{a \in \mathcal{A}} \mathbb{E}[r_t | a_t = a, \boldsymbol{\theta}].$$

3 | Problem Formulation

In this chapter, we formalize the problem we address in this thesis. We start in Section 3.1 by defining the setting we are considering and the related assumptions. Then, in Section 3.2, we present the learning problem and we define the metric we use for evaluating the quality of our solution.

3.1. Setting

We study the scenario in which we want to find the optimal pricing strategy for a set of products $\mathcal{J}$ (we call N the cardinality of set $\mathcal{J}$, i.e., $|\mathcal{J}| = N$). Our goal is, given a time horizon T, to set for every time $t \in \{1, \dots, T\}$ a vector of percentage margins (from now on, margins) $\mathbf{m}_t = (m_{1t}, \dots, m_{Nt})$ where $m_{jt} \in \mathcal{M}$ is the price we choose for product j at time t, and $\mathcal{M}$ is the (even infinite) set of possible margins. We define the (percentage) margin m_{jt} as:

$$m_{jt} := \frac{p_{jt} - c_j}{c_j},$$

where p_{jt} is the selling price and the acquisition cost for product j at time t, and c_j is its acquisition cost.

For a generic product $j \in \mathcal{J}$ we denote as $v_j(\mathbf{m}_t)$ the demand of product j which we assume to depend on the margin vector $\mathbf{m}_t$ of all products. We consider a scenario in which we may have both positive and negative interactions among the products, i.e., we assume that the purchase of a product has an effect on the purchases of other products. We identify two types of relationships between products: substitutability and complementarity. Consider two products $a, b \in \mathcal{J}$ and the vector margin $\mathbf{m}_t = (m_{1t}, \dots, m_{Nt})$. On the one hand, we call a and b substitutable products if the increases in the sales of one product imply a decrease in the sales of the other. On the other hand, we call a and b complementary products if the increases in the sales of one product imply an increment also in the sales of the other.

Assumptions We consider non-perishable products with unlimited availability, for which the demand function is monotonically non-increasing as its price increases. These assumptions, nowadays, hold in several cases. The assumption of unlimited availability virtually holds for e-commerce websites adopting the dropshipping (Singh et al., 2018) paradigm. The assumption of the monotonic demand holds for sellers vending products that are different from *Veblen*, *Giffen* or *Luxury* ones (Dougan, 1982; Leibenstein, 1950). We assume to be in a stationary environment.[1]

Available data We consider a scenario in which we assume to have access to transaction data reporting all the sales for every product $j \in \mathcal{J}$, divided by baskets. The only information we have about products is the groups of substitutable products. The detection of substitutable products has been thoroughly discussed in literature (Foxall et al., 2010; Van Gysel et al., 2018; Zhang et al., 2019) and is out of the scope of this work.

3.2. Learning Problem

The goal of our learning problem is to find the vector of the optimal margin $\mathbf{m}^*$, i.e., the vector maximizing our objective function $f(\mathbf{m})$. Formally:

$$\mathbf{m}^* \in \arg\max_{\mathbf{m} \in \mathcal{M}^N} f(\mathbf{m}), \tag{3.1}$$

where the objective function $f(\mathbf{m})$ is profit:

$$f(\mathbf{m}) := \sum_{j \in \mathcal{J}} m_j \; c_j \; v_j(\mathbf{m}), \tag{3.2}$$

over all the products.[2]

We call $\boldsymbol{\pi}$ a policy returning at each time t a vector of margins $\mathbf{m}_t$. We define the cumulative regret of such a policy as:

$$R(\boldsymbol{\pi}, T) := T f(\mathbf{m}^*) - \sum_{t=1}^{T} f(\mathbf{m}_t). \tag{3.3}$$

[1]The extension to non-stationary environments has been discussed several times (Bauer and Jannach, 2018; Nambiar et al., 2019; Javanmard et al., 2020; Mussi et al., 2022) and is out of the scope of this work.

[2]The objective function can be made more general by considering a convex combination of revenue and profit. In this case, the objective function f becomes $f(\mathbf{m}) := \sum_{j \in \mathcal{J}} (m_{jt} + 1 - \alpha) \; c_j \; v_j(\mathbf{m})$, with $\alpha \in [0, 1]$. In this formulation, it is easy to observe that if we select $\alpha = 1$ we optimize the profit as in Equation (3.2), and if we select $\alpha = 0$ we optimize the revenue.

The goal of our algorithm is to find a policy minimizing the expected cumulative regret:

$$\mathbb{E}\left[R(\boldsymbol{\pi}, T)\right] := Tf(\mathbf{m}^*) - \mathbb{E}\left[\sum_{t=1}^{T} f(\mathbf{m}_t)\right], \tag{3.4}$$

where the expectation is taken over the randomness of the realizations and the algorithm.

4 | Proposed Solution

In this chapter, we present our proposed algorithm to identify and price complementary products in an online manner. The algorithm consists of two parts. In the first part (Section 4.1), we deal with the aggregation of substitutable products and the discovery of complementary ones. In the second part (Section 4.2), we discuss the demand estimation model and the exploration strategy considered in the selection of optimal margins. In Figure 4.1 we show an outline of the proposed algorithm.

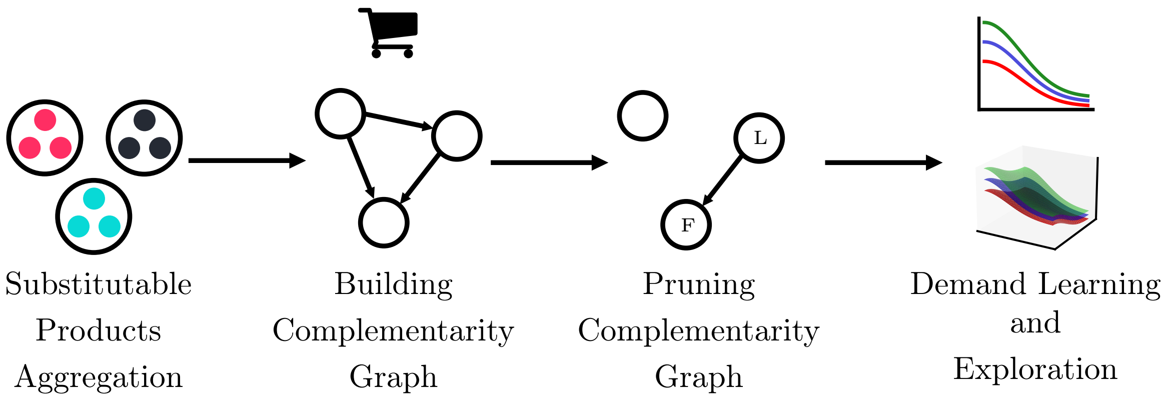


Figure 4.1: Algorithm outline

4.1. Complementary Products Discovery

Taking into account all the possible relations between every combination of products is challenging both in terms of computational and data requirements. A way to simplify the problem while focusing on the most significant relationships between products is necessary, in order to have an algorithm that scales properly with catalogues containing hundreds of thousands of products.

Conceptually, we can represent the relationships between products as a directed graph where each node is a product and each edge represents a complementarity relationship between 2 products. The graph is directed as the complementarity relationship is asymmetric, as is usually assumed in literature (Kocas et al., 2018; Yu et al., 2019). As stated

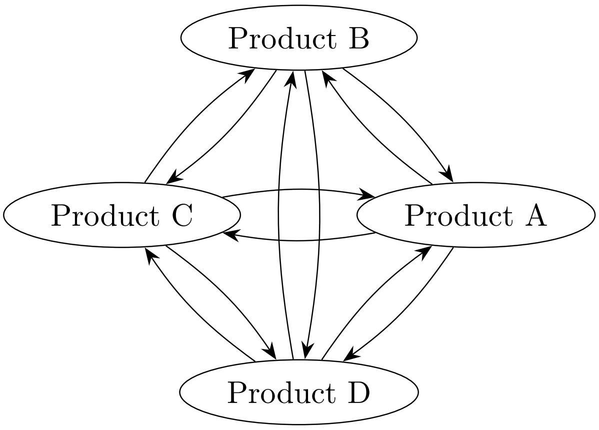


Figure 4.2: Example of a complementarity graph.

in Chapter 3, every product volume depends on the margins of the entire catalogue of products, so our initial representation of the graph of complementarity relationship (from now on the complementarity graph) is a connected graph, as shown in Figure 4.2.

In this section, we propose an approach to prune the edges of this graph while maintaining the most meaningful ones. The process makes use of information about substitutable products and transactional data and is divided into two steps:

1. Clustering substitutable products
2. Mining complementarity relationships.

4.1.1. Clustering Substitutable Products

The need to cluster substitutable products is due to the characteristics of these types of products. Two substitutable products compete with each other and are often subject to the phenomenon of cannibalization (Moorthy and Png, 1992) since they satisfy the same need. Applying dynamic pricing solutions to these products separately would exacerbate cannibalization and could lead to letting them compete against each other, with the effect of reducing profit. Clustering these products together allows them to be priced with the same pricing policy, minimizing cannibalization.

Following Mussi et al. (2022), given the information about substitutable products, which we assume to have, and a set $\mathcal{T}$ of time instants for which historical transaction data is available, we cluster a set of substitutable products $\mathcal{K} \subseteq \mathcal{J}$ and their historical margin $m_{k\tau}$ and volume $v_{k\tau}$ for all products $k \in \mathcal{K}$ and time $\tau \in \mathcal{T}$ in the following way:

$$m_{\mathcal{K}\tau} := \sum_{k\in\mathcal{K}} m_{k\tau} \cdot \frac{v_{k\tau}}{\sum_{h\in\mathcal{K}} v_{h\tau}},$$

$$v_{\mathcal{K}\tau} := \sum_{k\in\mathcal{K}} v_{k\tau}.$$

Given a time horizon T, the margin $m_{\mathcal{K}t}$ chosen at each time $t \in \{1, \ldots, T\}$ will be applied to every product $k \in \mathcal{K}$.

It should be noted that over time, the effect of averaging on margins, used at the beginning to aggregate historical data, becomes progressively more marginal, since the algorithm assigns, at each time $t \in \{1, \ldots, T\}$, the same margin for all aggregated products.

We refer to clustered substitutable products simply as products and each clustered product is a node in the complementarity graph. From now on we assume that there are no substitutable products in different nodes of the complementarity graph.

4.1.2. Mining Complementarity Relationships

After the identification of substitutable products, we proceed to identify meaningful complementary relationships between products.

As previously stated in Chapters 2 and 3, complementary products are frequently bought together and can then be identified using co-purchases in transactional data, as common in Market Basket Analysis (Ünvan, 2021). Following this approach, we propose a way to measure complementarity between products making use of the binomial test, discussed in Section2.2.2. We want to test the independence of every pair of products in order to verify if the co-occurrence of two products in the same basket is higher than *random chance* with a given significance level. More formally, given two products $a, b \in \mathcal{J}$, we denote the total number of baskets as n, the number of baskets containing product a as n_a, the number of baskets containing product b as n_b and the number of baskets containing both products a and b as n_{ab}.

The probability of having product a in a basket is:

$$P(a) = \frac{n_a}{n},$$

similarly for product b:

$$P(b) = \frac{n_b}{n}.$$

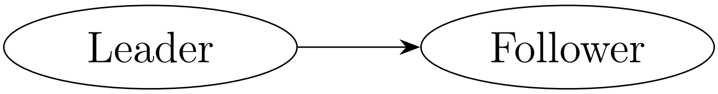


Figure 4.3: Example of a leader-follower relationship.

Under the assumption of independence, the probability of having both products a and b in a basket is $P(a)P(b)$ and this is the hypothesized probability of success π:

$$H_0 : \pi = P(a)P(b).$$

Since we want to test if the probability of the co-occurrence of a and b is higher than if they were independent, we perform a right-tailed test and our alternative hypothesis is:

$$H_1 : \pi > P(a)P(b).$$

The number of trials in our hypothesis test is the total number of baskets n, while the number of successes k is the number of co-occurrences of products a and b:

$$k = n_{ab}.$$

Finally, we want to perform the test with a significance level of 1%, so we reject H_0 if the p-value is smaller than 0.01. For the product pairs for which we reject the null hypothesis H_0, we have statistical evidence that the probability of products a and b appearing in the same basket is higher than if they were independent, and we can consider product a and b to be complementary. We set the direction of the complementarity relationship between two products to be from the product with the bigger number of baskets where it is bought to the one with less. We do so on the assumption that the product with more selling volumes has more influence on the product with less than vice versa, and in a pair we call the product that influences the other product *leader* while the product that is influenced *follower*, as shown in Figure 4.3. By imposing this direction we obtain a Weighted Directed Acyclic Graph (DAG) of complementarity relationships between products, where the p-values of the hypothesis tests are the weights of the edges. The absence of cycles is due to the unidirectional flow of the edges given by the product volumes.

In order to break down combinatorial complexity when optimizing we want to reduce the graph structure to multiple *star structures* composed of one leader and multiple followers,

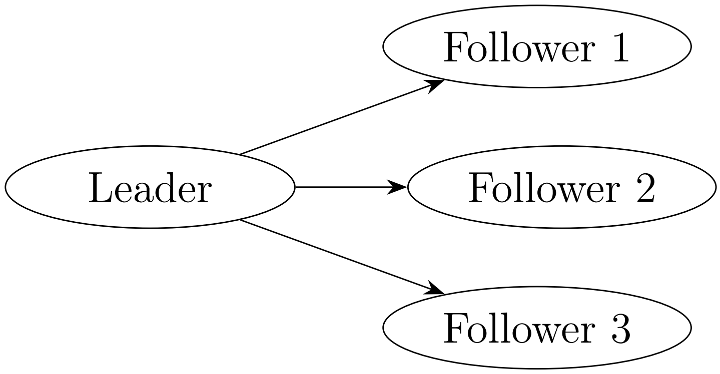


Figure 4.4: Star structure with one leader and multiple followers

where each node can only be either a leader or a follower, as shown in Figure 4.4.

In order to do this we first need to obtain tree structures from the DAG. This can be done by keeping, for each node, among the inbound edges, only the one with the highest weight, i.e. the most influential leader. It should be noted that we are guaranteed to obtain trees with this process thanks to the absence of cycles in the complementarity graph. Finally, we need to cut the trees into star structures. To accomplish this, we need to address the case where a node is both a leader and a follower of other nodes. We propose an algorithm to obtain the desired structure: starting from the leaves of the trees and going up to the roots, when we encounter a node that is both a follower and a leader, we only consider its leader role and, as a consequence, only keep its outbound edges. The result of this process is the above-mentioned structure: a collection of stars, with one leader and multiple followers. We have obtained a structure that allows us to perform optimization of the margins of the products breaking down the intrinsic combinatorial complexity of the problem. In Figure 4.5 we show a visual example of the graph pruning process, labelling with I the isolated nodes, L the leader nodes and F the follower nodes in the final graph.

4.2. Pricing Complementary Products

After mining meaningful complementarity relationships and reducing the complexity of the complementarity graph to obtain a structure feasible to optimize, in this section, we proceed to talk about the actual pricing methodology. We aim to find the margins that maximize the total profit (Equation 3.2). At each time t, our algorithm outputs the vector margins $\mathbf{m}_t$ and tackles the exploration-exploitation dilemma in order to minimize the expected regret (Equation 3.4).

In the following, we first discuss the demand estimation of products whose demand does not depend on other products, then we examine the demand estimation for products whose

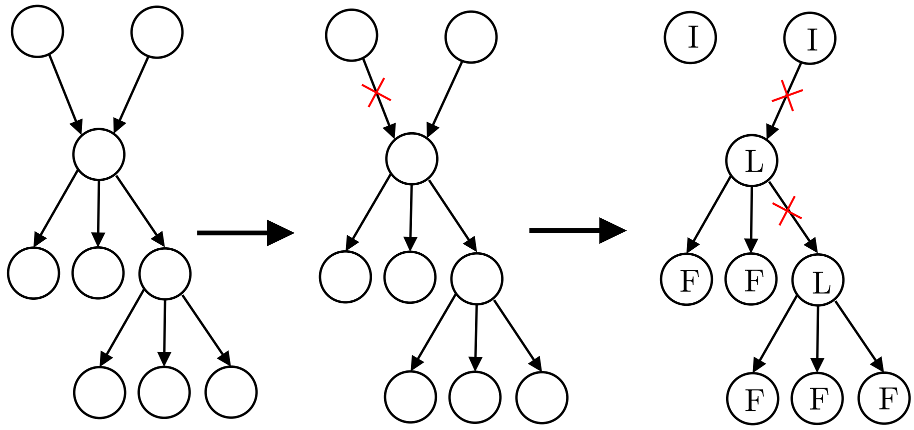


Figure 4.5: Example of graph pruning

demand depends on complementary products. Lastly, we address the chosen exploration strategy of the algorithm and the choice of the optimal vector margins $\mathbf{m}_t$ at time t.

4.2.1. Univariate demand learning

The resulting complementary graph after the steps discussed in Section 4.1 may have isolated nodes, which indicate products for which no complementary relations have been identified. Therefore, we model the demand for these products on the assumption that it depends only on their own margins.

Excluding the isolated nodes, the rest of the graph consists of the star structures above mentioned, where each edge represents a complementarity relationship. Similarly to the isolated nodes, we assume the demand of the leader products in the graph to depend only on their own margins. Given the set of products $\mathcal{J}$, we define the set of isolated products $\mathcal{I} \subseteq \mathcal{J}$ and for the star structures we define the set of leaders $\mathcal{L} \subset \mathcal{J}$ and the set of followers for each leader $i \in \mathcal{L}$ $\mathcal{F}_i \subset \mathcal{J}$, i.e. such that there is a directed edge $(i, j) \quad \forall j \in \mathcal{F}_i, \forall i \in \mathcal{L}$.

Formally, the demand of product $i \in \mathcal{I} \cup \mathcal{L}$ is such that:

$$v_i(\mathbf{m}) = v_i(m_i) \quad \forall \mathbf{m} \in \mathcal{M}^N,$$

where we denote with $v_i(\cdot)$ the demand of product i, with m_i the margin of product i and with N the total number of products. We estimate the demand using a Bayesian Linear Regressor (BLR, Tipping, 2001), this type of conditional model allows the estimation of uncertainty and will allow us to use a MAB approach to balance exploration and exploitation. Using the BLR we build an estimate $\hat{d}_i(\cdot)$ of the demand function for product i as a linear combination of the basis function taken as input, formally:

$$\hat{d}_i(m_i) = \sum_{h=0}^{Z} \theta_h \ \phi_h(m_i),$$

where θ_h is the h-th weight distribution and $\phi_h(m_i)$ is the h-th basis function of the margin $m_i \in \mathcal{M}$. In order to improve the robustness of our model we employ a shape-constrained model for which we impose the shape of the function to learn to be monotonically decreasing, following the assumption of monotonicity of demand made in Chapter 3. To do so we make use of the set of monotonically non-increasing basis functions discussed in Chapter 2 together with the choice of prior distributions with non-negative support.

4.2.2. Bivariate demand learning

While we assume the demand for the leader products in the graph to depend only on their own margins, we assume the demand for the follower products to depend on their own margin and on the margin of their leader.

Formally, given the set of leader products $\mathcal{L} \subset \mathcal{J}$ and the set $\mathcal{F}_i \subset \mathcal{J}$ of products that are follower of leader $i \in \mathcal{L}$, the demand functions of the follower products is such that:

$$v_j(\mathbf{m}) = v_j(m_j, m_i), \quad \forall j \in \mathcal{F}_i, \forall i \in \mathcal{L}, \forall \mathbf{m} \in \mathcal{M}^N.$$

Similarly to the univariate demand case, we employ a BLR to produce the estimate $\hat{d}_j(\cdot)$ of the demand function for product $j \in \mathcal{F}_i$, with $i \in \mathcal{L}$. For each follower product $j \in \mathcal{F}_i$ of leader product $i \in \mathcal{L}$, the demand is estimated as:

$$\hat{d}_j(m_j, m_i) = \sum_{h=0}^{Z} \theta_h \ \phi_h(m_j) + \sum_{h=0}^{Z} \theta_{h+Z+1} \ \phi_h(m_i),$$

where θ_h is the h-th weight distribution and $\phi_h(m)$ is the h-th basis function of the margin $m_j \in \mathcal{M}$. Following the characteristics of cross-price elasticity described in Section 2.1, we can also assume the additive contribution of leader products on follower volumes to be monotonically decreasing on their own margin. Thus, we employ the same set of monotonically non-increasing basis functions for both margins.

4.2.3. Exploration strategy and joint optimization

We can naturally frame our problem as an online learning one, where we want to acquire new information about the function we want to learn while at the same time minimizing the cumulative regret. The use of BLR allows us to measure uncertainty and use this information in a MAB setting to balance exploration and exploitation, where the arms of the MAB are the margins our algorithm chooses at each round. In order to do this we employ an approach similar to Thompson Sampling (TS, Agrawal and Goyal, 2012): in each round we draw a sample from each of the posterior distribution of the BLR weights, obtaining an estimation of the demand curve given the margins on which it depends. This is done both for isolated products and for products related to other products, obtaining a univariate estimation of the demand curve for isolated or leader products and a bivariate one for follower products. Given the estimation of the demand curve, we proceed to choose the margins that maximize the estimated profit. We distinguish the case where we price an isolated product and the one where we price related products.

For an isolated product $i \in \mathcal{I}$ we can compute the estimated objective function $\hat{f}_i(m), \forall m \in \mathcal{M}$ as:

$$\hat{f}_i(m) = m\ c_i\ \hat{d}_i(m), \quad \forall i \in \mathcal{I}.$$

We want to maximize $\hat{f}_i$, to do so we choose at each time $t \in \{1, \dots, T\}$ the optimal margin $\hat{m}^*_{it}$ for product i as:

$$\hat{m}^*_{it} = \arg\max_{m \in \mathcal{M}} \hat{f}_i(m).$$

With regard to the products belonging to a star structure, the estimated objective function for a leader product $i \in \mathcal{L}$ is, in the same way as the isolated products:

$$\hat{f}_i(m_i) = m_i\ c_i\ \hat{d}_i(m_i) \quad \forall i \in \mathcal{L},$$

while for the follower products $j \in \mathcal{F}_i$ of leader $i \in \mathcal{L}$ the estimated objective function is:

$$\hat{f}_j(m_j, m_i) = m_j\ c_j\ \hat{d}_j(m_j, m_i).$$

When choosing the optimal margins for the products belonging to a star structure, we jointly optimize the objective functions and obtain the vector margin $\hat{\mathbf{m}}_t$:

$$\hat{\mathbf{m}}_t = \underset{\mathbf{m}\in\mathcal{M}^N}{\arg\max} \sum_{i\in\mathcal{L}} \left[\hat{f}_i(\mathbf{m}) + \sum_{j\in\mathcal{F}_i} \hat{f}_j(\mathbf{m}) \right],$$

where N is the number of products belonging to $\mathcal{L} \cup \bigcup_{i\in\mathcal{L}} \mathcal{F}_i$.

5 | Related Works

This chapter explores the other works in the literature correlated to the one described in this thesis. We distinguish two main research fields related to our work: the one that deals with the identification of complementary products and the one that focuses on learning for dynamic pricing.

5.1. Complementary Products Identification

With regard to the identification of complementary products, the field of recommender systems has a rich literature on this subject, since discovering complementarity relationships among products can be used to recommend items complementary to the ones purchased by a user. Numerous works resort to the use of a product graph where the edges are complementarity relationships to represent the interactions between products. McAuley et al. (2015), using, as a primary source of data, product reviews and making use of topic modelling, builds a network of substitutable and complementary products. Sun et al. (2015) resort to association rules to identify implicit item-item relationships, accounting for the asymmetric nature of this relation. Zhao et al. (2017) builds a directed graph of complementary products using NLP techniques and in particular the Skip-Gram architecture to obtain product embeddings. Wan et al. (2018) uses a similar Skip-Gram architecture but integrates compatibility and loyalty estimates, while Wang et al. (2018) uses catalogue information together with transaction data to build a product graph of both substitutability and complementarity relationships and applies category and multi-step path constraints to the graph to discover new relationships. Hao et al. (2020) employs graph-based product representation learning using both customer behaviour and textual descriptions and predicts complementary product recommendations taking into account diversity, while Xu et al. (2020) uses co-view and product descriptions to build product knowledge graph embedding, using an architecture similar to Skip-Gram. Although these works demonstrate the effectiveness of representing the relationships between products in a graph, they serve a fundamentally different purpose from that of our work. In fact, unlike our setting, reducing the complexity of the graph is not a priority, as there is no need to jointly use and optimise bandit models. Moreover works like McAuley et al. (2015) make use of textual features when identifying complementary products, but they do not necessarily correspond to the purchasing dynamics of a specific e-commerce, based on which pricing can be done, unlike actual transactional data. Lastly, a further difference with our setting is the data requirements most of these works have, requiring textual descriptions, product reviews and co-views in addition to transactional data, the latter being the only information we assume to have to discover complementary products.

5.2. Learning for Dynamic Pricing

There is extensive literature on the estimation of demand and the consequent choice of optimal price given the demand curve estimated by the algorithm. Existing studies focus on this problem but neglect the crucial trade-off between exploration and exploitation, lacking theoretical guarantees. For example, Besbes and Zeevi (2015) enforces model-wise monotonicity and demonstrates the suitability of linear regression models for demand function modelling. Besbes and Zeevi (2009) and Broder and Rusmevichientong (2012) present parametric formulations of the demand function. Cope (2007) and Bauer and Jannach (2018) are seminal works on Bayesian inference applied to dynamic pricing, including even features such as competitors' prices, but they do not force monotonicity on the demand curve. In their work, Araman and Caldentey (2009) introduces a Bayesian framework that incorporates a monotonic formulation for the demand function. This formulation captures market-related information by integrating a prior belief on the parameters. Wang et al. (2021) explore non-parametric models for the estimation of a nonconcave demand function. Additionally, the authors make the assumption that the demand function exhibits smoothness and examine how this characteristic influences the robustness of the models. Shukla et al. (2019) propose a dynamic pricing algorithm that focuses on the monotonic willingness-to-pay of customers, making use of customer-related features that are missing in our setting. The aforementioned works focus on the dynamic pricing of single products, ignoring interactions between products.

Among the works that consider the relationships among complementary products, Mulhern and Leone (1991) is a seminal work on multiple-product pricing of retail stores, empirically demonstrating how a multiproduct approach boosts retail performance. Ghoniem et al. (2016) deals with a joint assortment and pricing problem on complementary retail categories, each category composed of substitutable products. Their definition of "primary" and "secondary" categories in a complementary relationship is a point of contact with our "leader-follower" framework, while the assumption of limited inventory differs from our setting. The problem is tackled through mixed-integer linear programming. Kocas et al. (2018) investigates the role of bestsellers on complementary products, modelling in addition traffic generation in online retailers in a duopoly scenario. Feng et al. (2018) models substitutability and complementarity through discrete choice models.

Unlike all the above-mentioned works, MABs provide theoretical guarantees, effectively tackling the exploration-exploitation dilemma. A landmark contribution which resorts to the MAB framework for dynamic pricing is Rothschild (1974). This approach has been extended in multiple directions. First, Kleinberg and Leighton (2003) tackles the problem of dealing with continuous-demand functions by proposing a discretization of the price values that provide theoretical guarantees on the algorithm's regret. Subsequently, Trovò et al. (2015, 2018) and Misra et al. (2019) exploit that, in many practical settings, the demand function is monotonically non-increasing in the price to design novel algorithms outperforming the classical MAB policies empirically. Mussi et al. (2022) applies a monotonic bandit to price long-tail products aggregating products with similar demands together. Mussi et al. (2023) considers the scenario in which the e-commerce website faces different customers and wants to find the optimal pricing scheme taking into account volume discounts. However, these approaches do not consider complementarity products.

To the best of our knowledge, the integration of complementary product identification, graph representation, and joint pricing through a bandit algorithm is unprecedented in the existing literature. This novel approach fills a gap in the current research landscape and provides a framework for addressing the intricate dynamics of complementary products in a pricing strategy that balances exploration and exploitation, while also being computationally feasible.

6 | Experimental Evaluation

In this chapter, we discuss the evaluation of our algorithm. The experimental evaluation of our algorithm is carried out through a synthetic environment with which we simulate the purchasing dynamics of customers and their price sensitivity. In Section 6.1 we describe the characteristics of our simulated environment while in Section 6.2 we compare `CPP` with pricing products independently.

6.1. Simulation Environment

We design a synthetic environment with the purpose of emulating market dynamics with reasonable assumptions. The environment is characterized by a fixed number of potential customers and a catalogue of products to buy.

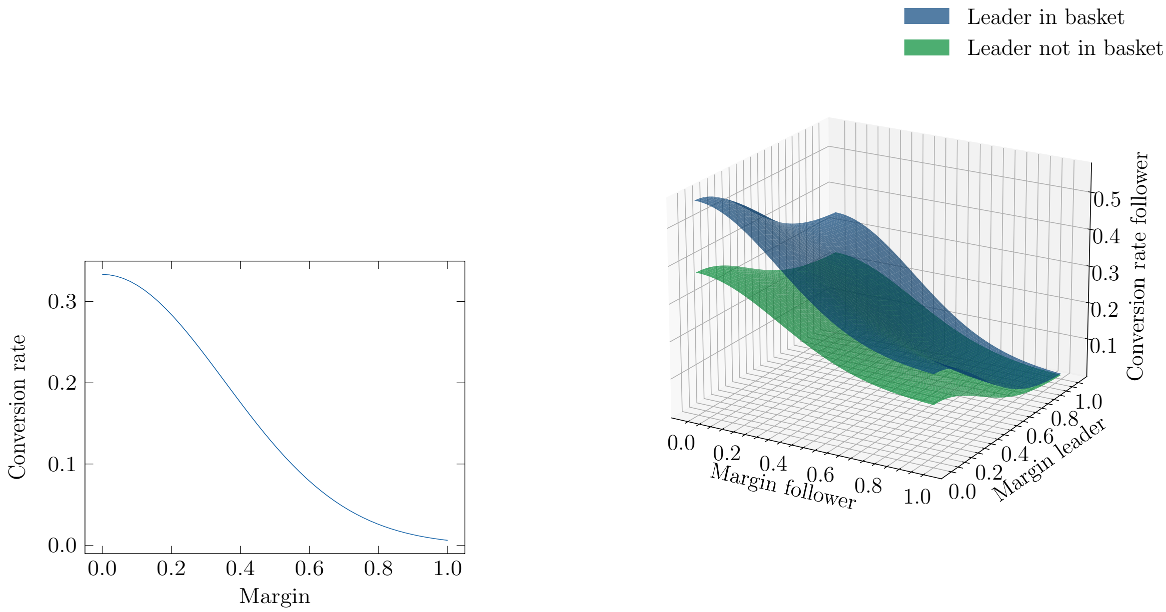
Each product is characterized by a *conversion rate*, the probability of purchasing such product given its margin. We denote the conversion rate of product $i \in \mathcal{J}$ given the vector margin $\mathbf{m} \in \mathcal{M}^N$ as $c_i(\mathbf{m})$.

Each product is related to others through fixed complementarity relationships. In the environment, we assign to each follower product one leader. We model the complementarity relationships between products by resorting to the conversion rates. Specifically, we model the conversion rate of follower products to be dependent on their own margins and on those of the leaders. In order to simulate the effect that leaders have on the purchase of followers in the same basket we model the conversion rate of followers to be conditioned on the purchase of their leader and to increase in case their leader is added to the basket. In particular, for a follower $j \in \mathcal{F}_i$ with $\mathcal{F}_i$ the set of followers of leader $i \in \mathcal{L}$, we scale the conversion rate by employing a coefficient l in the following way:

$$c_j(\mathbf{m}) \quad \text{when } i \text{ is not in the same basket}$$

$$l\ c_j(\mathbf{m}) \quad \text{when } i \text{ is in the same basket.}$$

The transactions of the users are generated by sampling a *Bernoulli* distribution for each



(a) Leaders (b) Followers

Figure 6.1: Conversion rates in synthetic environment.

product in the catalogue and using as parameter p the conversion rate of each product, such that:

$$P(B_i = 1) = c_i(\mathbf{m})$$

and

$$P(B_i = 0) = 1 - c_i(\mathbf{m}),$$

where B_i is the event of a purchase of product $i \in \mathcal{J}$ by a potential customer.

In Figure 6.1 we show the conversion rates used to generate transactions, univariate in the case of the leaders and bivariate on the margins of the followers of the leaders in the case of the followers. Specifically, the function used for the conversion rates is $f_1(x) = e^{-(2x)^2}/3$ for the leaders and $f_2(x, y) = f_1(x) + 0.7 f_1(y)$ for the followers. We limit the margins domain in $[0, 1]$. The conversion rate of the leader i is $c_i(m_i) = f_1(m_i)$ while for the follower j of leader i is $c_j(m_j, m_i)/l$ when there is not leader i in the same basket and $c_j(m_j, m_i)$ when leader i is present. Coefficient l is set to 1.5.

In Figure 6.2 we show, for a leader-follower pair, the expected profit function assuming

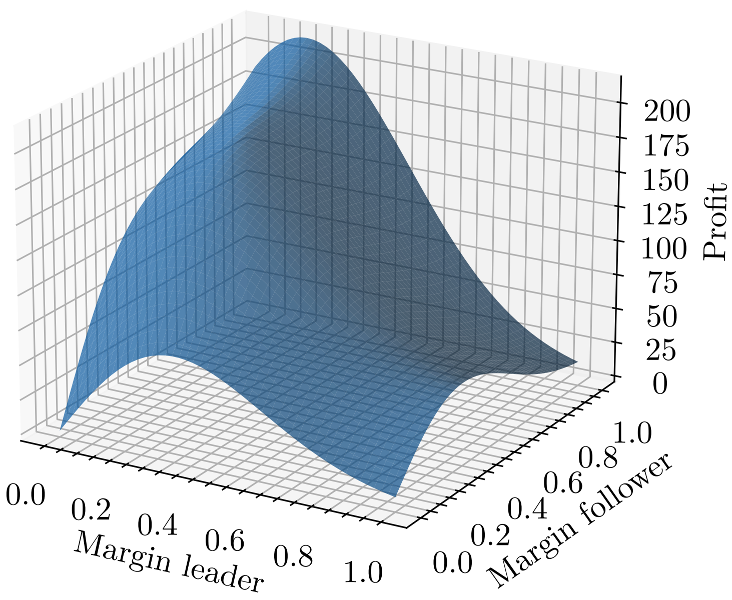


Figure 6.2: Expected profit of leader-follower pair, 1000 potential customers, cost 1.

1000 potential customers and the cost of both products to be 1.

We employ as basis functions in the demand estimators the monotonically non-increasing transformed Bernstein polynomial (see Section 2.3.4) of degree 10.

We employ as priors of the BLRs the Lognormal distribution with parameters $loc = 0.5, scale = 10$ for the bias and $loc = 0, scale = 10$ for the rest of the polynomials. We test our algorithm with a catalogue of 10 products with 5 leader-follower pairs and 1000 users.

At each timestep, the agent, using the transactional data observed up to that point, infers the complementarity relationships and fits the BLRs. It then chooses the optimal margins by managing exploration through Thompson Sampling. We split the margin domain into 100 arms that will be pulled by our agent. After pulling an arm, the agent observes the transactional data as a reward of the environment.

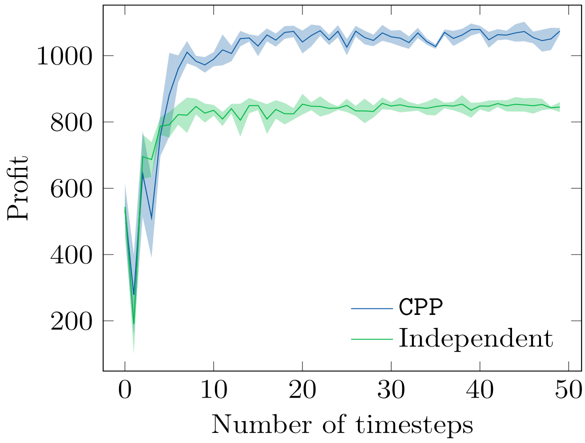


Figure 6.3: Instantaneous profits.

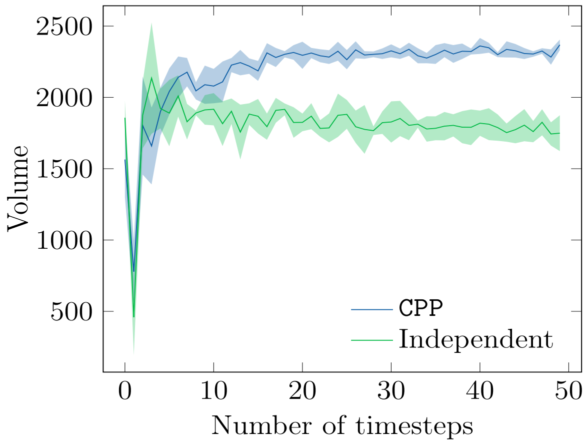


Figure 6.4: Instantaneous volumes.

6.2. Comparison with independently priced products

We compare the performance of CPP with an algorithm that prices the products independently with the same BLR employed for isolated and leader products. We do so to evaluate the profit increase obtained thanks to the joint optimization of complementary product prices.

In Figure 6.3 we show and compare the instantaneous profits obtained by jointly pricing complementary products with CPP and those obtained by independent pricing. We can observe that after 4 timesteps where the performance of the two approaches are comparable, CPP reaches a better optimum and unlocks profits up to 30% more w.r.t. independent pricing. As a side effect, we can observe in Figure 6.4 that, in this setting, choosing the margins with CPP results in an increase of volumes of over 30% as opposed to independent pricing.

7 | Conclusions

In this thesis, we faced the problem of finding the optimal pricing strategy for products presenting substitutable and complementary relations. We presented the problem under analysis, the related assumptions, and the learning problem, which consists of minimizing the expected regret. Then, we proposed `Complementary Product Pricing` (`CPP`), a novel strategy for learning online in this setting. The algorithm is composed of two main phases. In the former, we provided a strategy for the online identification of complementary relations. In the latter, we discussed a model for efficiently jointly optimizing the margin of the products. We conducted an extensive experimental campaign to assert the solution's soundness and goodness. The results showed that `CPP` effectively outperforms an independent pricing strategy, obtaining an increase of up to 30% in profits compared to independently priced products in a synthetic environment.

Future developments may consider removing the assumption of the knowledge of substitutable products. Another possible extension is dropping the non-stationary assumption on the environment, exploring the evolution of complementary relationships and demand of products over time. Finally, in order to identify complementary relationships we considered relations of products purchased in the same basket. An extension to this is to investigate the complementarity in purchases made over time by the same users.

Bibliography

Arnoud V. Den Boer. Dynamic pricing and learning: Historical origins, current research, and new directions. *Surveys in Operations Research and Management Science*, 20(1): 1–18, June 2015. ISSN 18767354. doi: 10.1016/j.sorms.2015.03.001. URL https://linkinghub.elsevier.com/retrieve/pii/S1876735415000021.

Ruiliang Yan and Subir Bandyopadhyay. The profit benefits of bundle pricing of complementary products. *Journal of Retailing and Consumer Services*, 18(4):355–361, July 2011. ISSN 09696989. doi: 10.1016/j.jretconser.2011.04.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0969698911000270.

R. Venkatesh and Wagner Kamakura. Optimal Bundling and Pricing under a Monopoly: Contrasting Complements and Substitutes from Independently Valued Products. *The Journal of Business*, 76(2):211–231, 2003. ISSN 0021-9398. doi: 10.1086/367748. URL https://www.jstor.org/stable/10.1086/367748. Publisher: The University of Chicago Press.

Walter Nicholson and Christopher Snyder. *Microeconomic theory: basic principles and extensions*. Cengage Learning, Australia ; Boston, MA, twelfth edition, 2017. ISBN 978-1-305-50579-7.

Julian McAuley, Rahul Pandey, and Jure Leskovec. Inferring Networks of Substitutable and Complementary Products. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 785–794, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 978-1-4503-3664-2. doi: 10.1145/2783258.2783381. URL https://doi.org/10.1145/2783258.2783381.

Francis J. Mulhern and Robert P. Leone. Implicit Price Bundling of Retail Products: A Multiproduct Approach to Maximizing Store Profitability. *Journal of Marketing*, 55(4): 63–76, October 1991. ISSN 0022-2429, 1547-7185. doi: 10.1177/002224299105500405. URL http://journals.sagepub.com/doi/10.1177/002224299105500405.

Ahmed Ghoniem, Bacel Maddah, and Ameera Ibrahim. Optimizing assortment and pricing of multiple retail categories with cross-selling. *Journal of Global Optimization*, 66(2):

291–309, October 2016. ISSN 0925-5001, 1573-2916. doi: 10.1007/s10898-014-0238-3. URL `http://link.springer.com/10.1007/s10898-014-0238-3`.

Guiyun Feng, Xiaobo Li, and Zizhuo Wang. On substitutability and complementarity in discrete choice models. *Operations Research Letters*, 46(1):141–146, January 2018. ISSN 01676377. doi: 10.1016/j.orl.2017.11.016. URL `https://linkinghub.elsevier.com/retrieve/pii/S0167637717301669`.

Cenk Kocas, Koen Pauwels, and Jonathan D. Bohlmann. Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling. *Journal of Interactive Marketing*, 41:28–43, February 2018. ISSN 10949968. doi: 10.1016/j.intmar.2017.09.001. URL `https://journals.sagepub.com/doi/full/10.1016/j.intmar.2017.09.001`.

Hang Yu, Lester Litchfield, Thomas Kernreiter, Seamus Jolly, and Kathryn Hempstalk. Complementary Recommendations: A Brief Survey. In *2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pages 73–78, May 2019. doi: 10.1109/HPBDIS.2019.8735479.

Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, and Kannan Achan. Product Knowledge Graph Embedding for E-commerce. In *Proceedings of the 13th International Conference on Web Search and Data Mining*, pages 672–680, January 2020. doi: 10.1145/3336191.3371778. URL `http://arxiv.org/abs/1911.12481`.

Simon Kemp. Perceiving luxury and necessity. *Journal of Economic Psychology*, 19(5): 591–606, October 1998. ISSN 0167-4870. doi: 10.1016/S0167-4870(98)00026-9. URL `https://www.sciencedirect.com/science/article/pii/S0167487098000269`.

William R. Dougan. Giffen Goods and the Law of Demand. *Journal of Political Economy*, 90(4):809–815, August 1982. ISSN 0022-3808. doi: 10.1086/261090. URL `https://www.journals.uchicago.edu/doi/10.1086/261090`. Publisher: The University of Chicago Press.

Sheldon M. Ross. *Introduction to Probability and Statistics for Engineers and Scientists*. Academic Press, 6 edition, 2021. ISBN 978-0-12-824346-6. URL `https://linkinghub.elsevier.com/retrieve/pii/B978012824346600003X`.

Christopher M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. Springer, New York, 2006. ISBN 978-0-387-31073-2.

S. McKay Curtis and Sujit K. Ghosh. A variable selection approach to monotonic regression with Bernstein polynomials. *Journal of Applied Statistics*, 38(5):961–976, May 2011. ISSN 0266-4763. doi: 10.1080/02664761003692423. URL `https:`

//doi.org/10.1080/02664761003692423. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02664761003692423.

Ander Wilson, Jessica Tryner, Christian L'Orange, and John Volckens. Bayesian non-parametric monotone regression. *Environmetrics*, 31(8):e2642, December 2020. ISSN 1180-4009, 1099-095X. doi: 10.1002/env.2642. URL https://onlinelibrary.wiley.com/doi/10.1002/env.2642.

Matthew D Hoffman and Andrew Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15 (1):1593–1623, 2014.

David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association*, 112(518): 859–877, April 2017. ISSN 0162-1459. doi: 10.1080/01621459.2017.1285773. URL https://doi.org/10.1080/01621459.2017.1285773. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.2017.1285773.

Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 1 edition, July 2020. ISBN 978-1-108-57140-1 978-1-108-48682-8. doi: 10.1017/9781108571401. URL https://www.cambridge.org/core/product/identifier/9781108571401/type/book.

Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret Analysis of Stochastic and Non-stochastic Multi-armed Bandit Problems, November 2012. URL http://arxiv.org/abs/1204.5721. arXiv:1204.5721 [cs, stat].

Shipra Agrawal and Navin Goyal. Analysis of Thompson Sampling for the Multi-armed Bandit Problem. In *Proceedings of the 25th Annual Conference on Learning Theory*, pages 39.1–39.26. JMLR Workshop and Conference Proceedings, June 2012. URL https://proceedings.mlr.press/v23/agrawal12.html. ISSN: 1938-7228.

Shipra Agrawal and Navin Goyal. Further Optimal Regret Bounds for Thompson Sampling. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, pages 99–107. PMLR, April 2013. URL https://proceedings.mlr.press/v31/agrawal13a.html. ISSN: 1938-7228.

Alessandro Nuara, Francesco Trovò, Dominic Crippa, Nicola Gatti, and Marcello Restelli. Driving Exploration by Maximum Distribution in Gaussian Process Bandits. *New Zealand*, 2020.

Gurpreet Singh, Harjot Kaur, and Amitpal Singh. Dropshipping in e-commerce: A perspective. In *Proceedings of the 2018 9th International Conference on E-Business, Management and Economics*, ICEME 2018, page 7–14, New York, NY, USA, 2018. Association for Computing Machinery.

H. Leibenstein. Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. *The Quarterly Journal of Economics*, 64(2):183–207, 1950. ISSN 0033-5533. doi: 10.2307/1882692. URL `https://www.jstor.org/stable/1882692`. Publisher: Oxford University Press.

Josef Bauer and Dietmar Jannach. Optimal pricing in e-commerce based on sparse and noisy data. *Decision Support Systems*, 106:53–63, February 2018. ISSN 0167-9236. doi: 10.1016/j.dss.2017.12.002. URL `https://www.sciencedirect.com/science/article/pii/S016792361730221X`.

Mila Nambiar, David Simchi-Levi, and He Wang. Dynamic Learning and Pricing with Model Misspecification. *Management Science*, 65(11):4980–5000, November 2019. ISSN 0025-1909. doi: 10.1287/mnsc.2018.3194. URL `https://pubsonline.informs.org/doi/10.1287/mnsc.2018.3194`.

Adel Javanmard, Hamid Nazerzadeh, and Simeng Shao. Multi-Product Dynamic Pricing in High-Dimensions with Heterogeneous Price Sensitivity, May 2020. URL `http://arxiv.org/abs/1901.01030`. arXiv:1901.01030 [cs, stat].

Marco Mussi, Gianmarco Genalti, Francesco Trovò, Alessandro Nuara, Nicola Gatti, and Marcello Restelli. Pricing the Long Tail by Explainable Product Aggregation and Monotonic Bandits. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3623–3633, Washington DC USA, August 2022. ACM. ISBN 978-1-4503-9385-0. doi: 10.1145/3534678.3539142. URL `https://dl.acm.org/doi/10.1145/3534678.3539142`.

Gordon R. Foxall, Victoria K. James, Jorge M. Oliveira-Castro, and Sarah Ribier. Product Substitutability and the Matching Law. *The Psychological Record*, 60(2):185–215, April 2010. ISSN 2163-3452. doi: 10.1007/BF03395703. URL `https://doi.org/10.1007/BF03395703`.

Christophe Van Gysel, Maarten de Rijke, and Evangelos Kanoulas. Mix 'n Match: Integrating Text Matching and Product Substitutability within Product Search. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, pages 1373–1382, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 978-1-4503-6014-2. doi: 10.1145/3269206.3271668. URL `https://dl.acm.org/doi/10.1145/3269206.3271668`.

Shijie Zhang, Hongzhi Yin, Qinyong Wang, Tong Chen, Hongxu Chen, and Nguyen Hung. *Inferring Substitutable Products with Deep Network Embedding.* August 2019. doi: 10.24963/ijcai.2019/598. Pages: 4312.

K. Sridhar Moorthy and I. P. L. Png. Market Segmentation, Cannibalization, and the Timing of Product Introductions. *Management Science*, 38(3):345–359, March 1992. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.38.3.345. URL `https://pubsonline.informs.org/doi/10.1287/mnsc.38.3.345`.

Yüksel Akay Ünvan. Market basket analysis with association rules. *Communications in Statistics - Theory and Methods*, 50(7):1615–1628, April 2021. ISSN 0361-0926. doi: 10.1080/03610926.2020.1716255. URL `https://doi.org/10.1080/03610926.2020.1716255`. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/03610926.2020.1716255.

Michael E. Tipping. Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*, 1(Jun):211–244, 2001. ISSN ISSN 1533-7928. URL `https://jmlr.csail.mit.edu/papers/v1/tipping01a.html`.

Zhu Sun, Guibing Guo, and Jie Zhang. Exploiting Implicit Item Relationships for Recommender Systems. In Francesco Ricci, Kalina Bontcheva, Owen Conlan, and Séamus Lawless, editors, *User Modeling, Adaptation and Personalization*, Lecture Notes in Computer Science, pages 252–264, Cham, 2015. Springer International Publishing. ISBN 978-3-319-20267-9. doi: 10.1007/978-3-319-20267-9_21.

Tong Zhao, Julian McAuley, Mengya Li, and Irwin King. Improving recommendation accuracy using networks of substitutable and complementary products. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 3649–3655, Anchorage, AK, USA, May 2017. IEEE. ISBN 978-1-5090-6182-2. doi: 10.1109/IJCNN.2017.7966315. URL `http://ieeexplore.ieee.org/document/7966315/`.

Mengting Wan, Di Wang, Jie Liu, Paul Bennett, and Julian McAuley. Representing and Recommending Shopping Baskets with Complementarity, Compatibility and Loyalty. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 1133–1142, Torino Italy, October 2018. ACM. ISBN 978-1-4503-6014-2. doi: 10.1145/3269206.3271786. URL `https://dl.acm.org/doi/10.1145/3269206.3271786`.

Zihan Wang, Ziheng Jiang, Zhaochun Ren, Jiliang Tang, and Dawei Yin. A Path-constrained Framework for Discriminating Substitutable and Complementary Products in E-commerce. In *Proceedings of the Eleventh ACM International Conference* *on Web Search and Data Mining*, pages 619–627, Marina Del Rey CA USA, February 2018. ACM. ISBN 978-1-4503-5581-0. doi: 10.1145/3159652.3159710. URL `https://dl.acm.org/doi/10.1145/3159652.3159710`.

Junheng Hao, Tong Zhao, Jin Li, Xin Luna Dong, Christos Faloutsos, Yizhou Sun, and Wei Wang. P-Companion: A Principled Framework for Diversified Complementary Product Recommendation. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 2517–2524, Virtual Event Ireland, October 2020. ACM. ISBN 978-1-4503-6859-9. doi: 10.1145/3340531.3412732. URL `https://dl.acm.org/doi/10.1145/3340531.3412732`.

Omar Besbes and Assaf Zeevi. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Science*, 61(4):723–739, 2015. Publisher: INFORMS.

Omar Besbes and Assaf Zeevi. Dynamic Pricing Without Knowing the Demand Function: Risk Bounds and Near-Optimal Algorithms. *Operations Research*, 57(6): 1407–1420, December 2009. ISSN 0030-364X. doi: 10.1287/opre.1080.0640. URL `https://pubsonline.informs.org/doi/abs/10.1287/opre.1080.0640`. Publisher: INFORMS.

Josef Broder and Paat Rusmevichientong. Dynamic Pricing Under a General Parametric Choice Model. *Operations Research*, 2012.

Eric Cope. Bayesian strategies for dynamic pricing in e-commerce. *Naval Research Logistics (NRL)*, 54(3):265–281, 2007. ISSN 1520-6750. doi: 10.1002/nav.20204. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.20204`.

Victor F. Araman and René Caldentey. Dynamic Pricing for Nonperishable Products with Demand Learning. *Operations Research*, 57(5):1169–1188, October 2009. ISSN 0030-364X. doi: 10.1287/opre.1090.0725. URL `https://pubsonline.informs.org/doi/abs/10.1287/opre.1090.0725`. Publisher: INFORMS.

Yining Wang, Boxiao Chen, and David Simchi-Levi. Multimodal Dynamic Pricing. *Management Science*, 67(10):6136–6152, October 2021. ISSN 0025-1909. doi: 10.1287/mnsc.2020.3819. URL `https://pubsonline.informs.org/doi/10.1287/mnsc.2020.3819`.

Naman Shukla, Arinbjörn Kolbeinsson, Ken Otwell, Lavanya Marla, and Kartik Yellepeddi. Dynamic Pricing for Airline Ancillaries with Customer Context. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, pages 2174–2182, New York, NY, USA, 2019. Association for

Computing Machinery. ISBN 978-1-4503-6201-6. doi: 10.1145/3292500.3330746. URL https://doi.org/10.1145/3292500.3330746.

Michael Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185–202, October 1974. ISSN 0022-0531. doi: 10.1016/0022-0531(74)90066-0. URL https://www.sciencedirect.com/science/article/pii/0022053174900660.

R. Kleinberg and T. Leighton. The value of knowing a demand curve: bounds on regret for online posted-price auctions. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 594–605, Cambridge, MA, USA, 2003. IEEE Computer. Soc. ISBN 978-0-7695-2040-7. doi: 10.1109/SFCS.2003.1238232. URL http://ieeexplore.ieee.org/document/1238232/.

Francesco Trovò, Stefano Paladino, Marcello Restelli, and Nicola Gatti. Multi–Armed Bandit for Pricing. *Proceedings of the 12th European Workshop on Reinforcement Learning*, pages 1–9, 2015.

Francesco Trovò, Stefano Paladino, Marcello Restelli, and Nicola Gatti. Improving multi-armed bandit algorithms in online pricing settings. *International Journal of Approximate Reasoning*, 98:196–235, July 2018. ISSN 0888-613X. doi: 10.1016/j.ijar.2018.04.006. URL https://www.sciencedirect.com/science/article/pii/S0888613X17303821.

Kanishka Misra, Eric M. Schwartz, and Jacob Abernethy. Dynamic Online Pricing with Incomplete Information Using Multiarmed Bandit Experiments. *Marketing Science*, 38(2):226–252, March 2019. ISSN 0732-2399. doi: 10.1287/mksc.2018.1129. URL https://pubsonline.informs.org/doi/10.1287/mksc.2018.1129. Publisher: INFORMS.

Marco Mussi, Gianmarco Genalti, Alessandro Nuara, Francesco Trovó, Marcello Restelli, and Nicola Gatti. Dynamic pricing with volume discounts in online settings. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 15560–15568, 2023.

List of Figures

List of Symbols

Symbol	**Meaning**
X	Constants
x	Variables
$\mathbf{x}$	Vectors
$\mathbf{M}$	Matrices
$\mathcal{A}$	Sets