OPTIMAL PRICING IN BLACK BOX PRODUCER-CONSUMER STACKELBERG GAMES USING REVEALED PREFERENCE FEEDBACK

Anup Aprem and Stephen J. Roberts

Department of Information Engineering, University of Oxford
{aaprem, sjrob}@robots.ox.ac.uk

ABSTRACT

This paper considers an optimal pricing problem for the black box producer-consumer Stackelberg game. A producer sets price over a set of goods to maximize his profit (the difference in revenue and cost function). The consumer buys quantity to maximize his utility function. The utility function of the consumer and the cost function of the producer are 'black box' functions (unknown functions with limited evaluations). Using a Gaussian process framework we derive a Bayesian algorithm for learning the optimal price. Numerical results illustrate the efficacy of our approach compared to existing literature.

***Index Terms*—** optimal pricing, Stackelberg game, producer-consumer, black box, Gaussian process

1. INTRODUCTION

Stackelberg games have been used to study sequential decision making in fields such as inventory and production control, wholesale and retail pricing strategies, outsourcing, and advertising. In a two-person Stackelberg game, the *leading* player (or leader) announces her action first and the other player, the *follower*, chooses an action to optimize his performance, given the policy of the leader. If the leader knows the objective function of the follower, then the leader anticipates the response of the follower and picks an action that optimizes her performance. The focus of this paper is on the case where the objective function of the leader and follower are 'black box' functions; the functions are unknown and the number of functional evaluations are limited by time or cost.

Context: In this paper, we consider the following problem (see Section 2.1 for details): A iterated game is played between a producer and a consumer. At each iteration, the producer decides the price p. In response to the price, the consumer buys x, which maximizes $v(x)-p^\intercal x$, where $v(\cdot)$ is the black box value function. The goal of the producer is to learn the unknown utility function of the consumer and set prices that maximize the profit $p^\intercal x - c(x)$, where $c(\cdot)$ is the black box cost function of the producer. The producer has access to the history of price and quantity purchased (the 'revealed preference' dataset) and the cost function evaluations to compute optimal price. Typical applications of the producer-consumer game include electricity pricing for smart-grid demand management and ad auctions; see [1] for a comprehensive list.

Related Work: The problem of estimating unknown utility function (or value function) can be traced back to area of revealed preference in microeconomics and more recently, in adversarial signal processing [2]. Major contributions to the area of revealed preference in microeconomics are due to Samuelson, Afriat [3] and Varian; see [4] for a recent survey. Given a revealed preference dataset, consisting of various price and quantity consumed by an agent, Afriat [3] devised a method to reconstruct the utility function. The revealed preference framework was later extended to games in [5]. Major drawbacks of the revealed preference approach of Afriat (and its extensions) is that there is no intelligent approach to select the optimal price and it cannot take into account the black box nature of cost and value function. Another line of work, in the context of games, is *Empirical game-theoretic analysis* [6, 7] where a dataset consists of a set of observations of strategy profiles and their corresponding payoffs (or utility values). However, this approach is different from our approach where the utility function is learnt only from strategies of players (or using revealed preference feedback).

More recent work exploits the concavity of the value function and cost function and uses gradient learning approach to compute the optimal price [8, 9]. Under these strong assumptions, the authors devise a polynomial time algorithm for computing optimal strategy of the leader. Whilst the concavity of the value function is well moti-

978-1-7281-0824-7/19/$31.00 ©2019 IEEE

vated by the rationality of the follower, modern production features such as indivisibilities, economies of scale and technology specialization make cost functions nonconvex [10]. Another major drawback is that the algorithms are not black box compatible; they are not *data-efficient*, i.e. the revealed preference data collected in one step of the algorithms is not re-used.

Main Results and Organization: Section 2.1 discusses the producer-consumer Stackelberg game and the key assumptions, and in Sec. 2.2, we discuss the Gaussian process framework, the framework adopted in this paper. Section 3 contains the main result: We derive a *data efficient* non-parametric Bayesian learning algorithm (Algorithm 1) to set optimal price for black box value and cost function. In addition, Algo. 1 makes no assumptions on the cost function. Experiments on synthetic examples are presented in Sec. 4 illustrating the efficacy of our approach. Finally, Sec. 5 offers concluding remarks.

2. PRELIMINARIES

2.1. Stackelberg Games

A iterated game is played between a producer and a consumer. At each iteration, the producer decides the price $p \in \mathbb{R}_+^d$, where d is the number of goods. In response to the price, the consumer buys (consumption bundle) $x^*(p) \in \mathbb{R}_+^d$, such that

$$x^*(p) = \underset{x \in \mathbb{R}_+^d}{\operatorname{argmax}} \left(v(x) - p^\top x \right), \tag{1}$$

where, $v(\cdot)$ is an unknown value function. The producer observes the purchased quantity. The profit of the producer is given by:

$$r(x) = p^\top x - c(x), \tag{2}$$

where, $c(\cdot)$ is the cost function. The objective of the producer is to set price that maximize his profit. The profit of the producer depends in turn on the unknown value function of the consumer. The producer relies on the prices and the quantity purchased (the revealed preference feedback) to learn about the unknown value function.

In this work, we use the following assumptions on the value function:

(A1) The value function is differentiable, monotone and concave.

(A2) The value function is homogeneous of degree χ.

Concave[1] and monotone[2] value functions (Assumption (A1)) are a consequence of human rational behavior. Assumption (A2) is satisfied by commonly used value functions such as the Cobb-Douglas function and Constant Elasticity of Substitution (CES) function.

Ass. (A1) and (A2) lead to the following theorems.

Theorem 1. *For any price vector p and optimal consumption x (obtained as a solution of* (1)*), $p = \nabla v(x)$*

The proof of Theorem 1 is straight forward from first order conditions and the fact the value function is concave (Ass. (A1)). Another interpretation, that we use in this paper, is that Theorem 1 says that the price gives the gradient of the value function at the optimal consumption bundle.

Theorem 2 (Euler's Theorem for homogeneous functions)**.** *Let $f : \mathbb{R}_+^d \to \mathbb{R}$, be a continuous and differentiable function. Then, f is homogeneous of degree χ, if and only if for all $x \in \mathbb{R}_+^d$, $x^\top \nabla f(x) = \chi f(x)$*

Theorem 3. *The profit of the producer is given by*

$$r(x) = \chi\, v(x) - c(x) \tag{3}$$

The proof of Theorem 3 is straight forward by substituting $p = \nabla v(x)$ (Theorem 1) in (2) and using the Euler's Theorem (Theorem 2). In comparison to (2), Theorem 3 expresses the profit function of the producer entirely in terms of the consumption bundle, as a function of the value function and the cost function.

2.2. The Gaussian Process (GP)

A Gaussian process f indexed by X is a stochastic process such that $\xi = (f(x_1), f(x_2), \cdots, f(x_n))$ is a Gaussian random vector for every $x_1, x_2, \cdots, x_n \in X$ [11]. The advantage of a Gaussian process is that it is completely specified by the mean and the covariance function (kernel) defined as follows:

$$\mu(x) = \mathbb{E}(f(x)),\ \kappa(x, \bar{x}) = \mathbb{E}\left((x - \mu(x))(\bar{x} - \mu(\bar{x}))^\top\right). \tag{4}$$

Hence, we denote the denote the Gaussian process by $f \sim \mathcal{GP}(\mu, \kappa)$.

Denote the dataset by $\mathcal{D} = \{x_k, f(x_k)\}_{k=1}^n$. Let, $m = (\mu(x_1), \mu(x_2), \cdots, \mu(x_n))$ and K be a $n \times n$ matrix such that $K_{i,j} = \kappa(x_i, x_j)$ then it can be easily seen from (4) that $\xi \sim \mathcal{N}(m, K)$. For any x^*, the predictive (or posterior) distribution is Gaussian, i.e. $f(x^*)|\mathcal{D} \sim \mathcal{N}(\mu^{\mathcal{D}}, (\sigma^{\mathcal{D}})^2)$ with mean and variance given by:

$$\mu^{\mathcal{D}} = K_*^\top K^{-1}(\xi - m),\ \ (\sigma^{\mathcal{D}})^2 = K_{**} - K_*^\top K^{-1} K_*, \tag{5}$$

[1] Concavity of value function models the human preference: *averages are better than the extremes*. It is also related to the law of diminishing marginal return, i.e. the rate of value decreases with x.

[2] Monotone value function models the human preference: *more is preferred to less*.

where $K_*(i) = \kappa(x_i, x^*)$ and $K_{**} = \kappa(x^*, x^*)$. An important property is that any affine transformation of a Gaussian process is also a Gaussian process [11].

3. LEARNING THE OPTIMAL STRATEGY IN STACKELBERG GAMES

In this section, we use the Gaussian process to model the black box cost and value function and learn the optimal price. In this paper, we consider the case $d = 1$[3].

3.1. System Model: GP for value function

The objective of the producer is to learn the unknown value and cost function, and set prices that maximize the profit function in (2). Since the profit function is expressed in terms of the value function and the prices as a gradient of the value function, we model the unknown gradient of the value function using a Gaussian process i.e.

$$\nabla v \sim \mathcal{GP}\left(\mu_{\nabla}(\cdot), \kappa_{\nabla}(\cdot,\cdot)\right). \tag{6}$$

The value function is obtained by integrating the gradient of value function. Since integration is a linear operation, the value function is a Gaussian process, and the joint Gaussian process between the value function and gradient is given by:

$$\begin{bmatrix} v(x) \\ \nabla v(x) \end{bmatrix} \sim \mathcal{GP}\left(\begin{bmatrix} \int \\ 1 \end{bmatrix} \mu_{\nabla}(x), \begin{bmatrix} \int \\ 1 \end{bmatrix} \kappa_{\nabla}(x,x') \begin{bmatrix} \int \\ 1 \end{bmatrix}^{\top}\right), \tag{7}$$

where $\int$ is used to represent the integration operation. The value function and the gradient should possess the structural properties in Ass. (A1) and (A2). To impose the monotone property of the value function, we require that the gradient of the value function be positive. We follow the approach in [12], and assume that the gradient function can be expressed as follows:

$$\nabla v(x) = \alpha + \frac{1}{2} h(x)^2, \tag{8}$$

where $h(x) \sim \mathcal{GP}\left(\mu_h(\cdot), \kappa_h(\cdot,\cdot)\right)$. In (8), α is a small positive constant. Note that the square transformation in (8) is non-linear, and hence, we follow the linearization approximation approach in [12]. The linearization approach involves approximating (8) using a first-order Taylor expansion, giving the following relation:

$$\begin{aligned} \mu_{\nabla}(x) &= \alpha + \frac{1}{2}\mu_h(x)^2, \\ \kappa_{\nabla}(x,x') &= \mu_h(x)\kappa_h(x,x')\mu_h(x'). \end{aligned} \tag{9}$$

[3]The extension to the general case is included in the forthcoming journal version.

Remark: In this work, we impose only the monotone assumptions on the Gaussian process model of the value function. The concavity constraints on the value function can be imposed by constraining the gradient of the value function (and its derivatives) as described in [13]. However, the resulting posterior distribution is not Gaussian distributed, and, obtaining moments (for subsequent steps) is complicated. The homogeneity constraint cannot be imposed due to the unknown constant of homogeneity. Not imposing the concavity and homogeneity constraints results in the algorithm being less efficient, albeit at a decreased computational cost.

3.2. Learning optimal strategy

In this section, we assume that the cost function and the degree of homogeneity χ is known. Section 3.3 discusses the case of an unknown cost function, and in Sec. 3.4, we will discuss how to infer the unknown χ.

We propose the following two step approach to compute the optimal price. In the first step, the producer proposes a consumption bundle. The proposal should balance the dual goals of exploring the unknown value function and exploiting the available information to reach the optimal price. In the second step, the producer sets a price to approximately induce the consumption bundle in the first step. The two steps are then iterated to arrive at an optimal price.

3.2.1. First Step: Proposing a consumption bundle

Let the revealed preference dataset be denoted as $\mathcal{D} = \left\{\left\{x^l, p^l\right\}_{l=1}^{k-1}\right\}$; the set of price and consumption bundles. The posterior distribution of the value function is a Gaussian process with mean and covariance function given by $\mu_v^{\mathcal{D}}(\cdot)$ and $\kappa_v^{\mathcal{D}}(\cdot,\cdot)$. The posterior distribution can be obtained using the dataset $\mathcal{D}$ and Eqns (5) to (9). Given the posterior distribution of the value function, the posterior distribution of the profit function is given as:

$$\mu_r^{\mathcal{D}}(x) = \chi\mu_v^{\mathcal{D}}(x) - c(x), \kappa_r^{\mathcal{D}}(x,x') = \chi^2\kappa_v^{\mathcal{D}}(x,x'), \tag{10}$$

where, we have assumed that the degree of homogeneity (χ) and the cost function ($c(\cdot)$) is known.

While choosing a consumption bundle, we need to balance the classical trade-off of exploring the unknown profit function and exploiting the available information. Hence, in this paper, we choose the Upper Confidence Bound (UCB) criterion in [14]:

$$\hat{x} = \underset{x}{\operatorname{argmax}} \left(\mu_r^{\mathcal{D}}(x) + \beta_k\sigma_r^{\mathcal{D}}(x)\right), \tag{11}$$

where, $\sigma_r^{\mathcal{D}}(x) = \sqrt{\kappa_r^{\mathcal{D}}(x,x)}$ and β_k is a parameter that controls the trade-off between exploring the unknown value function and exploiting the available data to set optimal price. The parameter β_k can be set using the annealing procedure in [14], or set to a constant.

3.2.2. Second Step: Inducing a consumption bundle

Given a proposed bundle, $\hat{x}$, a consumption bundle is induced through an appropriate choice of the price vector. The inducing price is given by the gradient of the value function at the proposed bundle $\hat{x}$. However, since the value function is modelled using a GP, we choose the price to be the expected value of the gradient of the value function (by integrating out the uncertainty). Hence,

$$p^k = \mathbb{E}_{\nabla}^{\mathcal{D}}(\hat{x}) = \mu_{\nabla}^{\mathcal{D}}(\hat{x}), \tag{12}$$

where, $\mu_{\nabla}^{\mathcal{D}}$ is the posterior distribution of the gradient obtained using the dataset $\mathcal{D}$, and Eqns (5) to (9). For price p^k, the consumer purchases quantity x^k.

3.3. Unknown cost function

In this section, we consider the case where a complete characterization of the cost function of the producer is not available. We show how to use the Gaussian process framework to model the cost function and hence, easily integrates with the framework in Sec. 3.2.

A surrogate model for the cost function can be built using a Gaussian process as follows: $c \sim \mathcal{GP}(\mu_c, \kappa_c)$, where μ_c and κ_c are the prior mean and covariance function of the cost function, respectively[4]. Given a dataset $\mathcal{D}_c = \{x_l, c(x_l)\}_{l=1}^{k}$, let the posterior mean and covariance function of the cost function be denoted by $\mu_c^{\mathcal{D}_c}$ and $\kappa_c^{\mathcal{D}_c}$, respectively. The posterior distribution can be obtained using the dataset $\mathcal{D}_c$ and (5). The posterior distribution of the profit function in (10) is then given by

$$\begin{aligned} \mu_r^{\mathcal{D}}(x) &= \chi \mu_v^{\mathcal{D}}(x) - \mu_c^{\mathcal{D}_c}(x), \\ \kappa_r^{\mathcal{D}}(x,x') &= \chi^2 \kappa_v^{\mathcal{D}}(x,x') + \kappa_c^{\mathcal{D}_c}(x,x'), \end{aligned} \tag{13}$$

where, χ is assumed to be known by the producer; see Sec. 3.4 for inferring the unknown χ. Note that while using (13) as the posterior distribution of the profit function, the UCB criterion in (11) also takes into account the uncertainty in the cost function.

[4] The formulation assumes that value function and cost function are independent. They can be modelled as a joint GP to reflect the combined relation of the quantity produced and its value function. However, here, it is assumed to be independent for notational simplicity.

Algorithm 1: Bayesian optimization algorithm to compute optimal price in Stackelberg games with unknown cost function and coefficient of homogeneity

Choose random initial price p^1.
Obtain initial quantity purchased x^1.
Dataset $\mathcal{D} = \{p^1, x^1\}$.
Cost function dataset $\mathcal{D}_c = \{x^1, c(x^1)\}$.
for $k = 2, \cdots$ **do**
    Estimate $\hat{\chi}_k$ using (15).
    Compute posterior mean $\mu_v^{\mathcal{D}}$ and covariance $\kappa_v^{\mathcal{D}}$ using the dataset $\mathcal{D}$ and Eqns (5) to (9).
    Compute posterior mean $\mu_c^{\mathcal{D}_c}$ and covariance $\kappa_c^{\mathcal{D}_c}$ of cost using dataset $\mathcal{D}_c$ and (5).
    Select $\hat{x}$ according to (11) and (13).
    Select price p^k according to (12).
    Obtain quantity purchased x^k for price p^k.
    Update dataset: $\mathcal{D} = \{\mathcal{D}, \{p^k, x^k\}\}$
    Update cost dataset $\mathcal{D}_c = \{\mathcal{D}_c \{x^k, c(x^k)\}\}$.

3.4. Inferring the coefficient of homogeneity

The coefficient of homogeneity χ can be inferred from the Gaussian process model of the value function, using Euler's Theorem (Theorem 2).

Given the dataset $\mathcal{D} = \{\{x^l, p^l\}_{l=1}^{k}\}$, let $\mathcal{V} = (v(x^1), v(x^2), \cdots, v(x^k))$ be the vector of (inferred) values at $x^1, \cdots, x^k$ respectively. Since the value function is a GP, $\mathcal{V}$ is a Gaussian random vector with mean $m_{\mathcal{V}} = (\mu_v^{\mathcal{D}}(x^1), \cdots, \mu_v^{\mathcal{D}}(x^k))$ and covariance matrix $K_{\mathcal{V}}(i,j) = \kappa_v^{\mathcal{D}}(x^i, x^j)$, where as in Sec. 3.2.1 we have denoted the posterior mean function and covariance function by $\mu_v^{\mathcal{D}}(\cdot)$ and $\kappa_v^{\mathcal{D}}(\cdot,\cdot)$, respectively.

Since, the value function is estimated from the integral of the gradient, the Euler's equation in Theorem 2 can be written as:

$$v(x^k) - \zeta = \frac{1}{\chi}(x^k)^\top p^k, \tag{14}$$

where, in (14) we have added a constant term (intercept) to include the effect of integration operation.

Let,

$$\Theta = \begin{pmatrix} (x^1)^\top p^1 & 1 \\ \cdot & \cdot \\ (x^k)^\top p^k & 1 \end{pmatrix}.$$

Then, Eqn (14) in vector notation is given by $\mathcal{V}$ =

$\Theta \begin{pmatrix} 1/\chi \\ \zeta \end{pmatrix}$. The generalized least-squares estimate of the degree of homogeneity and the intercept is given by

$$\begin{pmatrix} 1/\hat{\chi}_k \\ \hat{\zeta}_k \end{pmatrix} = \left(\Theta^{\intercal} K_{\mathcal{V}}^{-1} \Theta\right)^{-1} \Theta^{\intercal} K_{\mathcal{V}}^{-1} m_{\mathcal{V}}, \qquad (15)$$

where, the estimate of χ and ζ is indexed by k to indicate that it is estimated using the dataset $\mathcal{D}$.

Algo. 1 summarizes the various steps in learning optimal price in producer-consumer game with black box value and cost function.

4. NUMERICAL RESULTS

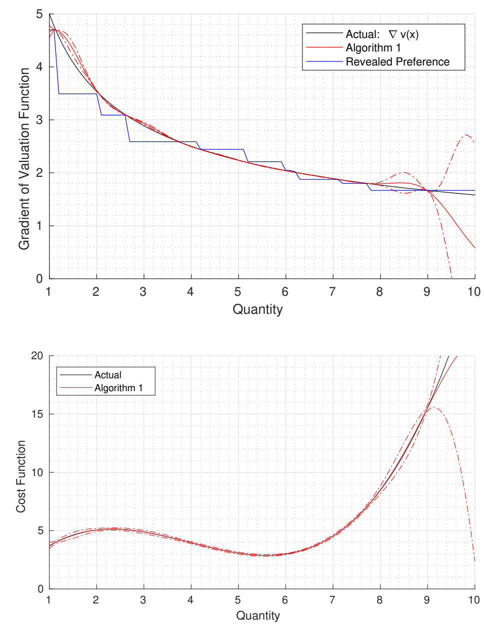
Consider the following example where the value function of the consumer is $v(x) = 10\sqrt{x}$ (the value function is monotone and concave) and the black box cost function of the producer is as shown in Fig. 1b. The cost function is adopted from [15], which describes the possibility that some firms face declining marginal costs over some ranges of output, due to internal labour market and capital utilization, leading to non-convexities in the cost function. Algo. 1 was run with the zero mean function and the squared exponential kernel given by:

$$\mu_h(x) = 0, \quad \kappa_h = \ell^2 \exp\left[-\frac{1}{2}\sum_{j=1}^{d} \frac{(x_j - \bar{x}_j)^2}{\lambda_j^2}\right], \qquad (16)$$

Setting the mean function to zero is motivated by the ordinal nature[5] of the value function. The choice of the squared exponential kernel is due to the ease of doing numerical integration in (7). The squared exponential kernel contains two hyperparameters: (i) the length scale parameter λ_j determines the relevance of each dimension, (ii) the parameter ℓ controls the magnitude of the output. The mean and kernel function of the gradient is given by (16) and (9). Similarly, the GP model for the cost function also uses a zero mean function and the squared exponential kernel. The hyperparameters were chosen by maximizing the log-likelihood; see Chap. 5 in [11].

Comparison with Revealed Preference [4]: As an illustration, Fig. 1a shows the actual gradient of the value function (in black), $\nabla v = 5/\sqrt{x}$, superimposed with the mean value of the Gaussian process model $\mu^{\mathcal{D}}_{\nabla v}$ (shown in red). The red dotted lines show the 95% confidence bounds ($\mu^{\mathcal{D}}_{\nabla v} \pm 2\sigma^{\mathcal{D}}_{\nabla v}$). The posterior distribution of the Gaussian process captures the structure of the value function which is not surprising since it has a simple monotone structure. The figure also shows (in blue) the gradient of the value function inferred using the revealed preference approach in microeconomics. The revealed preference approach produces a value function that is approximated using linear piecewise segments, and hence the 'staircase' nature of the gradient. However, note that the revealed preference approach is an offline approach and doesn't provide a criterion to intelligently select optimal price. In addition, the revealed preference approach cannot handle the black box cost function.

[5]An ordinal utility function is unique up to increasing monotone transformation. For example, adding a positive constant to the utility function does not change the solution in (1).



(b) Cost function

Fig. 1: Illustration of learning the value function of the consumer and cost function of the producer in the Stackelberg game in (1). Fig. 1a shows the actual gradient (in black), $\nabla v = 5/\sqrt{x}$, the mean value of the Gaussian process model $\mu^{\mathcal{D}}_{\nabla v}$ (shown in red), the gradient inferred using the revealed preference approach in microeconomics (in blue). In Fig. 1b, we show the posterior distribution of the cost function ($\mu_c^{\mathcal{D}_c}$ in red against the actual cost function (shown in black). The red dotted lines show the 95% confidence bounds.

In Fig. 1b, we show the posterior distribution of the cost function ($\mu_c^{\mathcal{D}_c}$ in red and $\mu_c^{\mathcal{D}_c} \pm 2\sigma_c^{\mathcal{D}_c}$ in dotted lines) against the actual cost function (shown in black). As can be seen from Fig. 1b, the posterior distribution captures the non-convex nature of the cost function. Capturing this non-convex nature makes our approach unique compared to existing literature.

Comparison with gradient based approach: [9] also use a two step approach as in Sec. 3.2. The first step on finding a proposal consumption uses a global optimizer similar to our approach. The second step on inducing the consumption bundle, uses a gradient approach, the gradient approach only uses the properties of value function (stronger versions of Ass. (A1) to (A2)). The strong convexity of cost function assumed in [9] ensures that the difference in consumption (difference between the proposed consumption in first step and the actual consumption in second step) can be bounded. This technique cannot be used for non-convex functions such as in Fig. 1b. Moreover, inducing the consumption bundle (second step) for a nominal tolerance of 0.1 uses an average of around $1e4$ gradient steps (refer to Learn-Price algorithm in [9]). Algo 1 (terminated when the change in price is less than $1e{-}2$), averaged over 100 independent simulations, is able to converge to the optimal price with 21 ± 2.58 iterations. In addition, the computed homogeneity parameter is equal to 0.49 (the actual value is 0.5), which corresponds to the homogeneity of the unknown value function chosen above. Hence, our approach is efficient for black box consumer-producer games i.e. when the functional evaluations (either the value or cost function) is limited.

General d: The authors expect, from existing results in [14], Algo. 1 to scale quadratically with respect to d compared to quintic scaling in [9].

5. CONCLUSION

We considered the problem of optimal price setting for producer-consumer games in the context of black box value and cost function. Algo. 1 details a Bayesian algorithm to compute the optimal price, using weaker assumptions compared to existing literature. Numerical results illustrate the efficiency compared to existing methods. The GP framework adopted in this paper is quite flexible and can accommodate the heterogeneity in the preference of consumer and time variation in the value or cost function[6].

6. REFERENCES

[1] Hal R Varian, "Revealed preference and its applications," *The Economic Journal*, vol. 122, no. 560, pp. 332–338, 2012.

[2] V. Krishnamurthy and W. Hoiles, "Afriat's test for detecting malicious agents," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 801–804, 2012.

[3] S. Afriat, "The construction of utility functions from expenditure data," *International economic review*, vol. 8, no. 1, pp. 67–77, 1967.

[4] Hal R Varian, "Revealed preference," *Samuelsonian economics and the twenty-first century*, pp. 99–115, 2006.

[5] Rahul Deb, "A testable model of consumption with externalities," *Journal of Economic Theory*, vol. 144, no. 4, pp. 1804–1816, 2009.

[6] Michael P. Wellman, "Methods for empirical game-theoretic analysis," in *Proc. of AAAI*, 2006, pp. 1552–1555.

[7] Patrick R. Jordan, L. Julian Schvartzman, and Michael P. Wellman, "Strategy exploration in empirical games," in *Proc. of AAMAS*, 2010, pp. 1131–1138.

[8] Kareem Amin, Rachel Cummings, Lili Dworkin, Michael Kearns, and Aaron Roth, "Online learning and profit maximization from revealed preferences," in *Proc. of AAAI*, 2015.

[9] Aaron Roth, Jonathan Ullman, and Zhiwei Steven Wu, "Watch and learn: Optimizing from revealed preferences feedback," in *Proc. of STOC*, 2016, pp. 949–962.

[10] Kristiaan Kerstens and Shunsuke Managi, "Total factor productivity growth and convergence in the petroleum industry: Empirical analysis testing for convexity," *Int. Journal of Production Economics*, vol. 139, no. 1, pp. 196 – 206, 2012.

[11] Carl Edward Rasmussen, *Gaussian processes in machine learning*, Springer, 2004.

[12] Tom Gunter, Michael A. Osborne, Roman Garnett, Philipp Hennig, and Stephen J. Roberts, "Sampling for inference in probabilistic models with fast Bayesian quadrature," in *Proc. of NIPS*, 2014, pp. 2789–2797.

[13] X. Wang and J. Berger, "Estimating shape constrained functions using Gaussian processes," *SIAM Journal on Uncertainty Quantification*, vol. 4, no. 1, pp. 1–25, 2016.

[14] Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger, "Gaussian process optimization in the bandit setting: no regret and experimental design," in *Proc. of ICML*, 2010, pp. 1015–1022.

[15] Valerie A. Ramey, "Nonconvex costs and the behavior of inventories," *Journal of Political Economy*, vol. 99, no. 2, pp. 306–334, 1991.

[6] These extensions are included in the forthcoming journal version.