Research Article

Warranty Optimization in a Dynamic Environment

Nedialko B. Dimitrov[1] and Stefanka Chukova[2]

[1] *Department of Mechanical Engineering, University of Texas at Austin, 1 University Station C2200, Austin, TX 78712, USA*

[2] *School of Mathematics, Victoria University of Wellington, P.O. Box 600, Wellington 6140, New Zealand*

Correspondence should be addressed to Stefanka Chukova, schukova@mcs.vuw.ac.nz

A product warranty is an agreement offered by a producer to a consumer to replace or repair a faulty item, or to partially or fully reimburse the consumer in the event of a failure. Warranties are very widespread and serve many purposes, including protection for producer, seller, and consumer. They are used as signals of quality and as elements of marketing strategies. In this study we review the notion of an online convex optimization algorithm and its variations, and apply it in warranty context. We introduce a class of profit functions, which are functions of warranty, and use it to formulate the problem of maximizing the company's profit over time as an online convex optimization problem. We use this formulation to present an approach to setting the warranty based on an online algorithm with low regret. Under a dynamic environment, this algorithm provides a warranty strategy for the company that maximises its profit over time.

1. Introduction

A product warranty is an agreement offered by a producer to a consumer to replace or repair a faulty item, or to partially or fully reimburse the consumer in the event of a failure. Warranties are very widespread and serve many purposes, including protection for producer, seller, and consumer. They are used as signals of quality and as elements of marketing strategies. A general treatment of warranty analysis is given by Blischke and Murthy [1, 2].

From the buyer's point of view, the main role of a warranty in any business transaction is protectional. Specifically, the warranty assures the buyer that faulty item will either be repaired or replaced at no cost or at a reduced cost. A second role of warranty is informational, as it implicitly sends out a message regarding the quality of the product and could influence buyer's purchase decision.

The main role of warranty from the producer's point of view is also protectional. Warranty terms may and often do specify the use and conditions of use for which the product

is intended and provide for limited coverage or no coverage at all in the event of misuse of the product. A second important purpose of warranty for the seller is promotional. As buyers often infer a more reliable product when a long warranty is offered, this has been used as an effective advertising tool. In addition, warranty has become an instrument, similar to product performance and price, used in competition with other manufacturers in the marketplace.

Despite the fact that warranties are so commonly used, the study of warranties in many situations remains an open problem. This may seem surprising since the fulfillment of warranty claims may cost companies large amounts of money. Underestimating true warranty costs will result in losses for a company, overestimating them will result in uncompetitive product prices. The data relevant to the modeling of warranty costs in a particular industry are usually highly confidential, since they are commercially sensitive. Much warranty analysis therefore takes place in internal research divisions in large companies.

The common warranty parameters of interest analyzed and evaluated are the expected warranty cost and the expected warranty cost per unit time over the warranty length for a particular item as well as the life cycle of the product; see Chukova and Hayakawa [3, 4]. Typically, the warranty length and the warranty policy are assumed to be known, which identifies the failure model. Based on the adopted model of the failure process, the total expected total warranty cost and sometimes the variance of this cost are evaluated.

The study presented here deviates from the traditional framework in warranty analysis. For simplicity we assume that the warranty is one-dimensional and nonrenewing, that is, the warranty is identified by its length, and it starts at the time the item is sold or the service has began. We consider time periods, such that the manufacturer's profit functions, as functions of warranty, and the warranty may vary from time period to time period. In general, we assume that the optimal warranty and the profit functions are unknown, but the profit for the assigned warranty in any particular time period is known. The aim of this study is to present an approach that will assure that if the warranty varies in a particular way, suggested by an online algorithm, under reasonable, quite general assumptions on the profit functions, the long run average of the manufacturer's profit will be comparable with profit if the optimal warranty was known at the time the product was launched on the market.

The outline of this paper is as follows. In Section 2 we present a brief overview of the online algorithm approach. The profit model is introduced in Section 3, and it is analysed, using an online algorithm, in Section 4. Section 5 contains concluding remarks.

2. Miscellaneous Results: Online Algorithm

In this paper, we concern ourselves with profit maximization, thus we consider the online convex programming problem with a sequence of concave functions and a maximization objective. In its simplest form, an online convex programming problem $(F, \{p_1, p_2, \ldots\})$ consists of a feasible region $F \in \Re^n$ and an infinite sequence of concave "profit" functions $\{p_1, p_2, \ldots\}$, each going from F to $\Re$. An algorithm for the online convex programming problem $\mathcal{A}(F, \{p_1, p_2, \ldots\})$ is an algorithm that produces a point w_i, which is a function only of the points $w_1, w_2, \ldots, w_{i-1}$, each previously produced by the algorithm, and the first $(i-1)$ functions $p_1, \ldots, p_{i-1}$. The regret of an algorithm is defined as

$$R(T) = \left(\max_{w^* \in F} \sum_{i=1}^{T} p_i(w^*) \right) - \sum_{i=1}^{T} p_i(w_i). \tag{2.1}$$

Interpreting, the regret measures the performance of the algorithm, which does not know p_i before producing w_i, to pick the single best point w^* in the feasible region F given knowledge of all the p_i's in advance.

Online convex optimization, introduced by Zinkevich [5], was originally motivated by the notion of playing repeated games. Imagine playing an infinitely repeating game that proceeds in rounds. In round i we must pick a strategy knowing only the strategies we have chosen in the previous rounds, from 1 to $(i-1)$, and the payoffs we received in those rounds. That is the motivation of the algorithm, which produces the point w_i knowing only $w_1, w_2, \ldots, w_{i-1}$ and the first $(i-1)$ functions $p_1, \ldots, p_{i-1}$. Each w_i can be thought of as the strategy in the ith round, and each p_i can be thought of as the payoff function in the ith round. The payoff function may change from round to round arbitrarily, since we do not know the strategies adopted by opponents in the game. In the repeated game settings, the regret then measures the amount of utility lost by a player who follows the strategy as specified by the algorithm versus picking the single best strategy to follow in all rounds.

Zinkevich exhibits an algorithm $\mathcal{A}_f(F, \{p_1, p_2, \ldots\})$, in full-information settings (see the Appendices), with regret $R(T) = O(\sqrt{T})$, which gives

$$\lim_{T \to \infty} \frac{R(T)}{T} \leq 0. \tag{2.2}$$

Interpreting, in the limit, following the strategies specified by the algorithm produces the same per period profit as picking the optimal single strategy. The quantity $R(T)/T$ is commonly referred to as the average regret.

Online convex optimization has clear industrial applications. For example, consider a company producing a product. The company's profit could be a concave function of the warranty offered by the company. However, the profit does not only depend on the warranty, but it could also depend on the types of products offered by competitors or the changing demands of customers. The profit function of the company in period i could be thought of as the p_i in the online convex optimization problem, and the w_i could be the warranty offered by the company in period i. An algorithm with low regret gives a warranty strategy for the company to follow that maximizes the company's profit over time.

One of the main hurdles to applying Zinkevich's algorithm directly is that it requires full knowledge of the function p_i after round i. In specific, Zinkevich's algorithm uses the gradient of the function p_i. However, in realistic settings, such as the example in the previous paragraph, a company may not know the entire function p_i. Instead, all the company learns in round i is the value of $p_i(w_i)$. In other words, all the company learns is the amount of profit the company made in round i, not the entire profit function. Flaxman et al. [6] exhibit an algorithm for online convex optimization, $\mathcal{A}_b(F, \{p_1, p_2, \ldots\})$, in bandit settings (see the Appendices), using only the value $p_i(w_i)$ of the profit function of the previous round and with regret $R(T) = O(T^{3/4})$.

Another concern with the direct application of online convex optimization is that the average regret results are in the limit as the number of rounds goes to infinity. Traditional industries, such as car manufacturing, have warranty on the order of years. Thus, even a few periods of the repeated profit maximization may take a human lifetime. However, warranties come in many varieties, and today's markets can be largely autonomous. For example, consider a competition between online brokerage firms. A firm could offer a warranty on the amount of time required to execute a purchase or sell an order. The warranty offered could change dynamically throughout the trading day. The broker's customers could themselves be

automated programs that dynamically choose which brokerage firm to use to execute trades. In such a scenario it is easy to imagine thousands of profit maximization rounds per day. Regardless of the plausibility of using online convex optimization in a specific application, the average regret results imply the startling conclusion that a company can attain nearly maximum profit in a dynamically changing environment, without knowing anything about the future.

In this paper, we study online convex optimization as applied to the warranty applications described in this section.

3. The Profit Model

In what follows, we propose a general form of profit functions $\{p_1, p_2, \ldots\}$ to be used with the two online convex optimization algorithms $\mathcal{A}_f$ and $\mathcal{A}_b$ for the warranty optimization examples in Section 2. Firstly, similarly to Bell et al. [7], we define the market share function $m(w)$ as a function of warranty w as follows

$$m(w) = \frac{a + gw}{a + gw + c}, \tag{3.1}$$

where a is a parameter of initial "attractiveness" or "reputation" of the company, g is the increase of the total attractiveness $(a + gw)$ of the company per unit increase of warranty, and c is the total attractiveness of the competitors of the company in the marketplace. It is easy to see that $m(w)$ is an increasing function of w.

This form of the profit function is appropriate in modeling different market structures. For example, if $c = 0$, the company has a monopoly in the marketplace, whereas altering the value of c will model the arrival or departure of a competitor.

To gain some intuition on the market share function, suppose that the warranty w is zero. We then have

$$m(0) = \frac{a}{a + c}. \tag{3.2}$$

One can think of this equation as follows. Suppose a customer picks which company to use randomly, but with weights proportional to the company's attractiveness. The form of $m(0)$ in (3.2) is the probability the customer selects to do business with our company instead of a competitor, given that the company assigns no warranty to its products. Another interpretation of (3.2) is that, if the company assigns no warranty to its products it will have $m(0)$ share of the market. Now, consider form (3.1) and let $w \to \infty$. We have $\lim_{w \to \infty} m(w) = 1$, which means that if the company offers a large warranty it will dominate the entire market.

Now, using the market share function $m(w)$ given in (3.1), we introduce the profit function $p(w)$, again as a function of warranty. We propose

$$p(w) = \mathbf{P}m(w) - \mathbf{R}m(w)F(w), \tag{3.3}$$

where $\mathbf{P}$ is a constant equal to the total market value of the considered industry, $\mathbf{R}$ is a constant equal to the penalty of total recall of all sold products, and $F(w)$ is the cumulative distribution function of the lifetime X of the product. The latter function represents the quality and

reliability of the production and governs the process of failures and related warranty claims. We assume a linear relationship between **P** and **R** of the following form:

$$\mathbf{R} = \gamma \mathbf{P}, \quad 1.0 \leq \gamma \leq 2.0. \tag{3.4}$$

In the case of $\gamma \leq 1$, even if all products are recalled, offering a large warranty will guarantee that the company will end up with a profit. On the other hand, if $2.0 \leq \gamma$, in order to avoid heavy penalties, the most appropriate strategy for the company is to sell the product with no warranty. Therefore, in both of these cases the optimal strategy of the company is known, and we will focus our study on the nontrivial case of $1.0 \leq \gamma \leq 2.0$.

4. Modeling a Dynamic Environment

In what follows we display the performance of $\mathcal{A}_f$ and $\mathcal{A}_b$ in several differing models of a dynamic environment. First, we present an environment with a quality improvement under two failure scenarios: a gradual failure modeled with an exponential lifetime distribution and a shock failure modeled with a Weibull lifetime distribution. Second, we present an environment with increasing competition again under two scenarios: a gradual increase in competition and a shock increase in competition. Finally, we present an environment where we increase the penalty for faulty products. We show that in all these environments, the algorithms $\mathcal{A}_f$, and $\mathcal{A}_b$ perform well as compared to algorithm "opt_fixed" which selects a single, optimal warranty for all rounds, even though neither $\mathcal{A}_f$ nor $\mathcal{A}_b$ know the future profit functions. As algorithm $\mathcal{A}_b$ is a randomized algorithm and its theoretical guarantees are in expectation, in each scenario we present the expected behavior of $\mathcal{A}_b$ over 50 independent runs. In addition, we include in the comparison the algorithm "opt_round" that selects the optimal warranty in each round.

4.1. Environment with Quality Improvement

Refer to the profit functions defined in (3.3). Our next goal is to use these functions for decision making related to warranty, in environment with quality improvement. We model the dynamic environment with quality improvement by using the cumulative distribution function $F_X(w)$ of the lifetime X of the product. We consider two cases.

Case 1. Firstly, we assume that $X \sim \text{Exp}(\lambda)$, that is,

$$F_X(w) = 1 - e^{-\lambda w}, \tag{4.1}$$

and the mean time to failure is $E(X) = 1/\lambda$. Based on (4.1) we define a sequence of profit functions as

$$p_i(w) = \mathbf{P}m(w) - \mathbf{R}m(w)\left(1 - e^{-\lambda_i w}\right), \tag{4.2}$$

and use them in full-information settings, that is, with $\mathcal{A}_f(F, \{p_1, p_2, \ldots\})$ as well as in bandit settings, that is, with $\mathcal{A}_b(F, \{p_1, p_2, \ldots\})$. The results are presented in Figure 1.
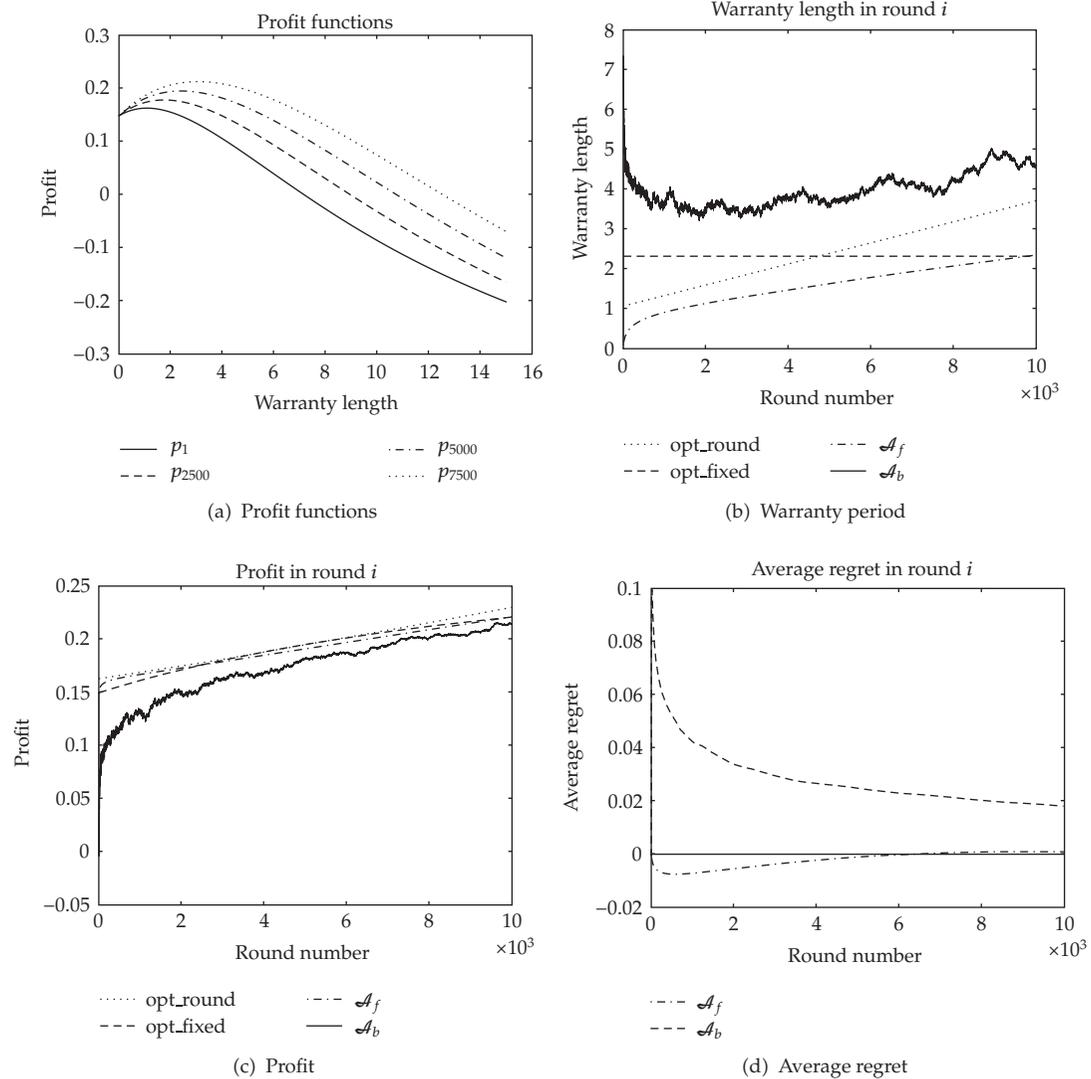
(a) Profit functions

(b) Warranty period

(c) Profit

(d) Average regret

Figure 1: In (a) are represented several profit functions, with p_i representing the profit function in round i. In (b) is represented the warranty offered, as a function of the round number, by various algorithms. (c) represents the profit earned, as a function of the round number, by each algorithm. Finally, (d) represents the average regret of the two algorithms, $\mathcal{A}_f$ and $\mathcal{A}_b$. The algorithm "opt_fixed" selects a single, optimal warranty for all rounds. The algorithm "opt_round" selects the optimal warranty in each round.

In this example, we model quality improvement by additively increasing the parameter of the exponential distribution representing the lifetime of the product. The mean of the distribution changes linearly from 4 to 8. The resulting profit functions are presented in Figure 1(a). As you can see, in later rounds, as the quality increases, the company can offer a larger warranty to capture a larger fraction of the market and thus receive higher profit. Figure 1(b) shows the warranty offered by the various algorithms. Algorithm $\mathcal{A}_f$ starts by offering a zero warranty, the imposed initial starting point, and follows an upward trend as the rounds increase. Algorithm $\mathcal{A}_b$ has an initial starting point, dictated by the algorithm

itself, around the middle of the feasible region. In all our examples, the feasible region is $\{w \mid 0 \leq w \leq 15\}$, that is, the acceptable warranty is between 0 and 15. That is why, initially, the warranty of $\mathcal{A}_b$ decreases from 7.5 and then increases as the rounds increase. Figure 1(c) shows the profit earned in each round by each algorithm. The figure illustrates the benefits of using $\mathcal{A}_b$, as it closely follows the profit received by optimizing the warranty in each round, but it assumes very limited information of the profit functions. Figure 1(d) shows how the average regret of $\mathcal{A}_b$ decreases to zero as the rounds increase. In other words, the per period loss of $\mathcal{A}_b$ as compared to following the optimal fixed warranty decreases to zero as the rounds increase. Even better results are pictured in Figure 1(d) for $\mathcal{A}_f$; however, it assumes knowledge of the gradient of the profit function in each round, where as $\mathcal{A}_b$ only assumes knowledge of the evaluation of the profit at a single point.

Case 2. Secondly, we assume that $X \sim \text{Weibull}(\gamma)$, that is,

$$F_X(w) = 1 - e^{-w^\gamma}, \tag{4.3}$$

and the mean time to failure is $E(X) = \Gamma(1 + 1/\gamma)$ and create the sequence of profit functions

$$p_i(w) = \mathbf{P}m(w) - \mathbf{R}m(w)\big(1 - e^{-w_i^\gamma}\big). \tag{4.4}$$

In this example, we introduce quality improvement with a Weibull lifetime distribution. In a Weibull distribution, there is a sharp threshold at which most products fail. That is why in Figure 2(a) the profit functions fall sharply as the warranty increases. Figure 2(c) represents the profit of the various algorithms. Notice that the profit for "opt_fixed" begins negatively and sharply increases as the rounds increase. This is because the single warranty chosen by "opt_fixed" in early rounds is greater than the failure threshold of the product, but is less than the failure threshold in later rounds. The profit earned by $\mathcal{A}_b$ is negative in early rounds, since $\mathcal{A}_b$ begins with an initial point in the middle of the feasible region, which is much larger than the failure threshold of the initial Weibull distributions. As the rounds increase, $\mathcal{A}_b$ decreases the warranty, as pictured in Figure 2(b). Since both the failure threshold increases and $\mathcal{A}_b$ decreases the warranty in later rounds, $\mathcal{A}_b$ eventually begins to make a profit. In late rounds, $\mathcal{A}_b$ begins to approach the performance of $\mathcal{A}_f$ and "opt_round", which outperform "opt_fixed", since they can increase the warranty as the failure threshold increases. As expected, since $\mathcal{A}_f$ outperforms "opt_fixed", the average regret for $\mathcal{A}_f$, pictured in Figure 2(d), is negative. The average regret for $\mathcal{A}_b$ increases in the early rounds, while $\mathcal{A}_b$ is making poor profit, and quickly decreases in the later rounds.

4.2. Environment with Increasing Competition

We model the increase in competition in the profit function through the parameter c included in the market share function (3.1). In this example, we additively increase the competition from 2 to 50, with the parameter a set to 1. Interpreting, this means that initially the company has roughly 1 to 2 odds of attracting a customer. Toward the final round, the company has only 1 to 50 odds of attracting a customer, thus the competition has increased. Figure 3(b), through the graph for "opt_round", shows that the warranty that should be offered by the company increases as competition increases. This is to capture a larger fraction of the market
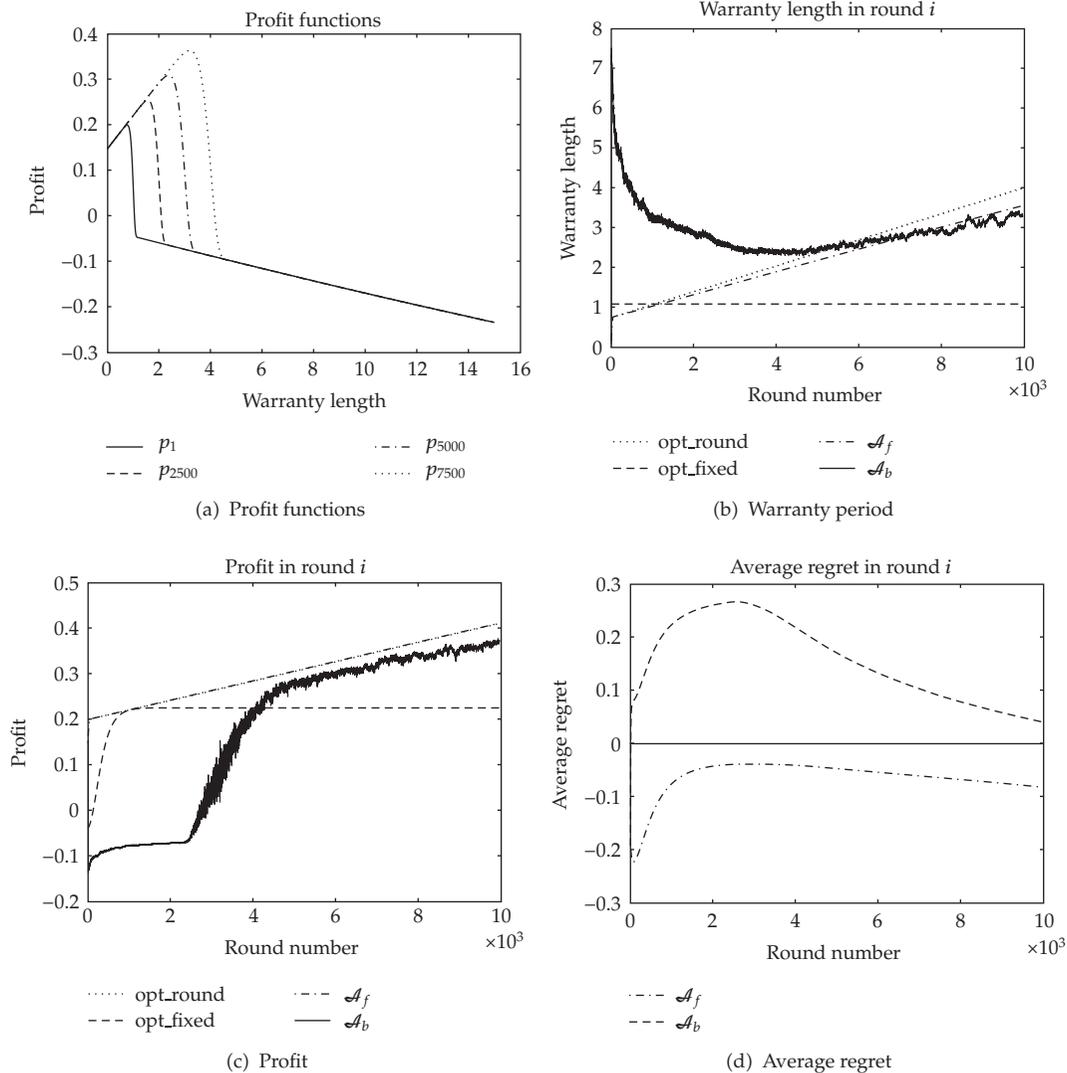
(a) Profit functions

(b) Warranty period

(c) Profit

(d) Average regret

Figure 2: These figures represent the algorithms' performance with a Weibull lifetime distribution. In a Weibull lifetime distribution, there is a sharp threshold before which most products are functioning properly and after which most products have failed. That is why in (a) we see profit functions that fall sharply as the warranty increases.

as dictated by expression (3.1). The warranty of $\mathcal{A}_b$ decreases throughout, as it initially begins at the middle of the feasible region. Algorithm $\mathcal{A}_f$, on the other hand, begins initially with a warranty of zero and closely follows the performance of "opt_round". Figure 3(d) shows that $\mathcal{A}_b$ looses less than 15% of the total profit at the end of the example. This percentage would decrease to zero as the rounds go to infinity, by the results of Flaxman et al. [6].

This example shows the algorithm behavior when there is a shock increase in competition. In round 2000, the arrival of a competitor decreases the market share of the company significantly; the value of c jumps from 2 to 50. This leads to the sudden drop pictured in Figure 4(c), for all warranty settings, even for "opt_round". Though the profit for
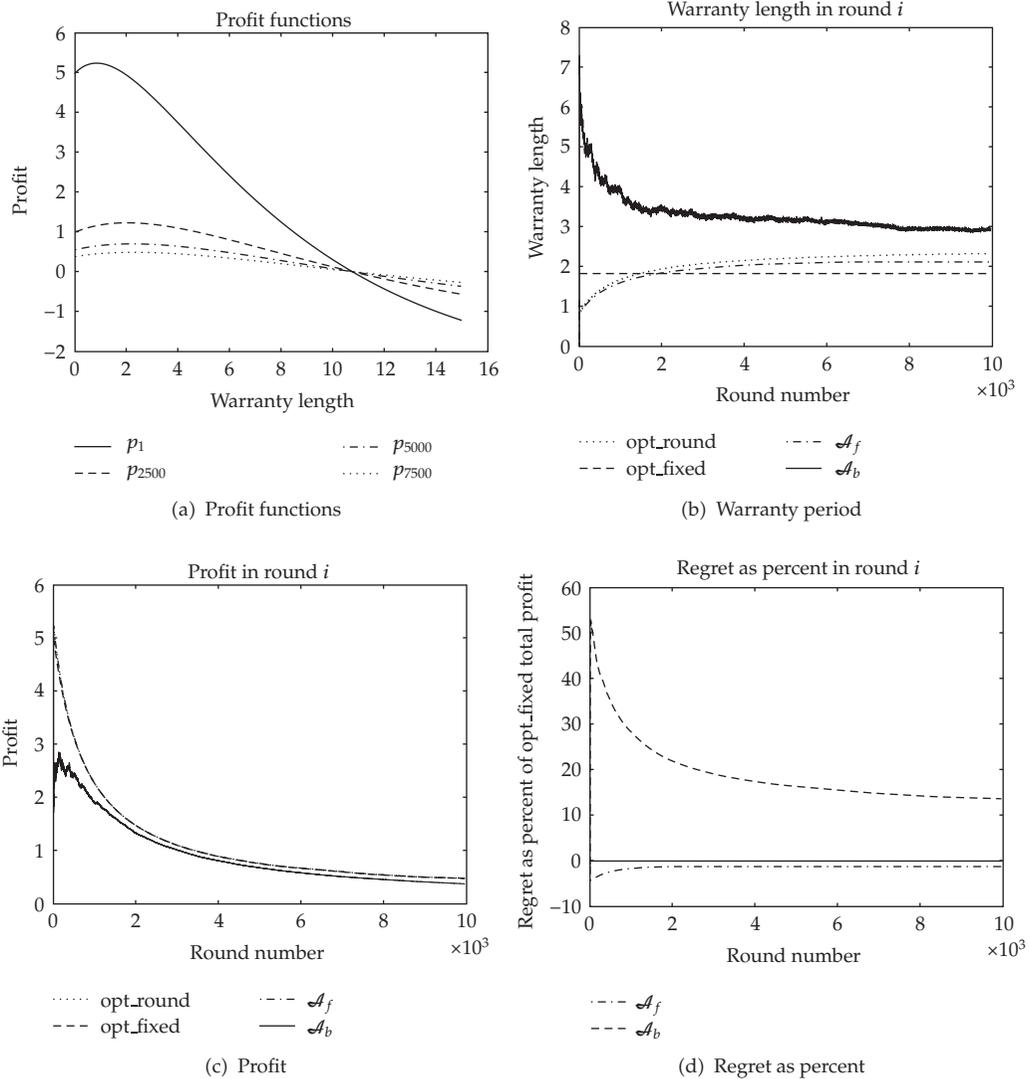
(a) Profit functions

(b) Warranty period

(c) Profit

(d) Regret as percent

Figure 3: In this example, the competition increases additively from 2 to 50. As the competition increases, the company's share of the market decreases and so does the profit as shown in (a). (d), shows the regret in round i as a percentage of the total profit received by "opt_fixed" from rounds 0 through i. It illustrates that the theoretical results showing that the average regret tends to zero as the rounds tend to infinity translate into results showing that the percent of profit missed by $\mathcal{A}_b$ and $\mathcal{A}_f$ tends to zero as the round tends to infinity.

all algorithms has a sudden drop, it is interesting to see the algorithm's reaction in changing the warranty, pictured in Figure 4(b). Again, due to the different information settings of $\mathcal{A}_f$ and $\mathcal{A}_b$, algorithm $\mathcal{A}_f$ is near the optimal setting before the competition increase and needs a short time to readjust after the increase. On the other hand, $\mathcal{A}_b$ begins, as usual, in the middle of the feasible region and is decreasing the warranty toward the optimal setting in the initial rounds. After the competition increase, $\mathcal{A}_b$ continues to decrease the warranty but at a slower pace. Even though the warranty offered by $\mathcal{A}_b$ seems far from the warranty offered by the
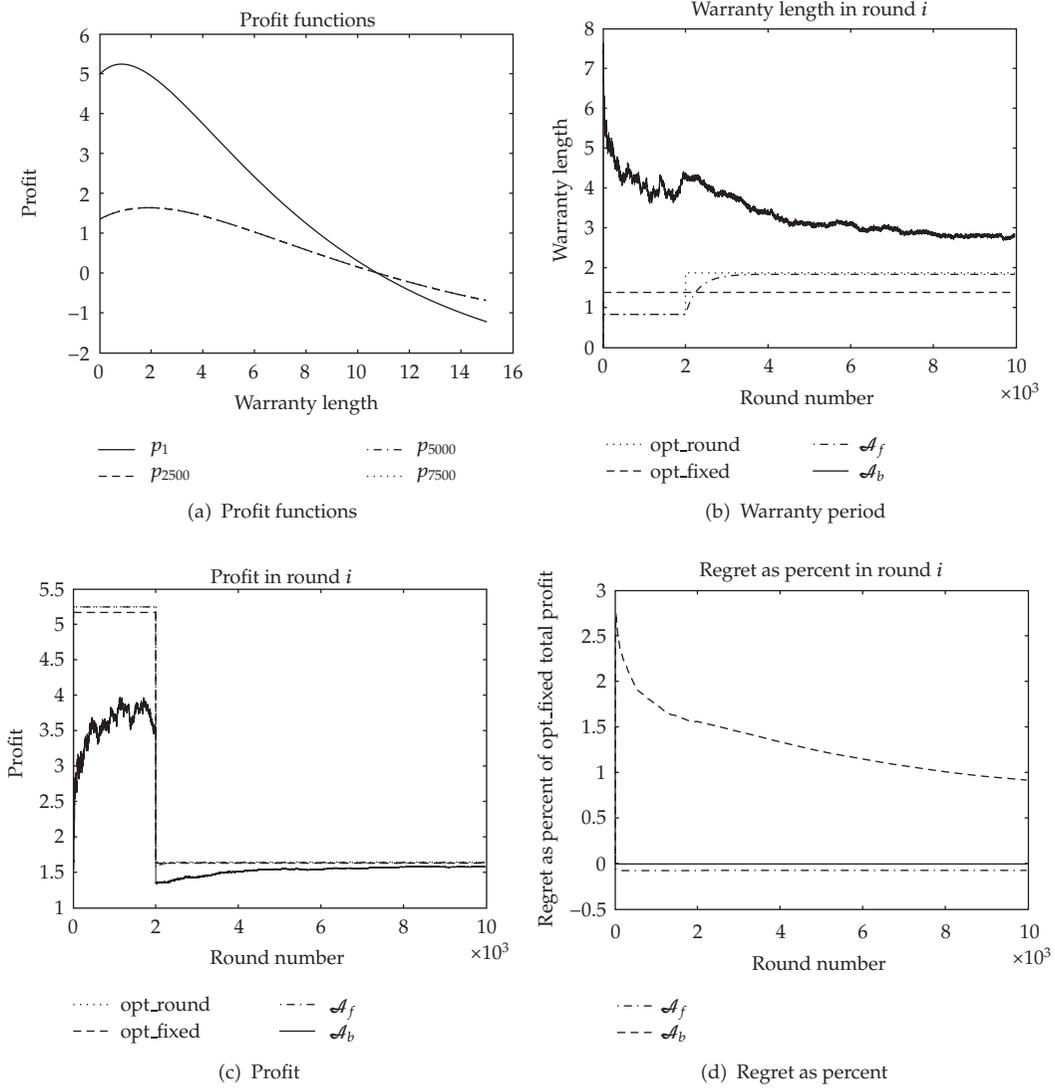
(a) Profit functions

(b) Warranty period

(c) Profit

(d) Regret as percent

Figure 4: These graphs represent a shock increase in competition. In round 2000, the parameter c in the market share increases from 2 to 50.

other algorithms, Figure 4(d) shows that its regret, as a percentage of the total profit gained by "opt_fixed", is once again steadily decreasing toward zero.

4.3. Environment with Changeable Penalties

In this example, we study a linear increase in the penalty from a faulty product. In specific, we alter the ratio γ between **P** and **R** in the profit function (3.3) from 1 to 2. A larger γ models a larger cost to replace a failed item. As can be seen in Figure 5(a), as the penalty for a faulty product increases, the optimal warranty goes to zero. Figure 5(b) shows how the warranty of $\mathscr{A}_f$ starts at zero, increases until it passes the optimal warranty offered by "opt_round",

(a) Profit functions

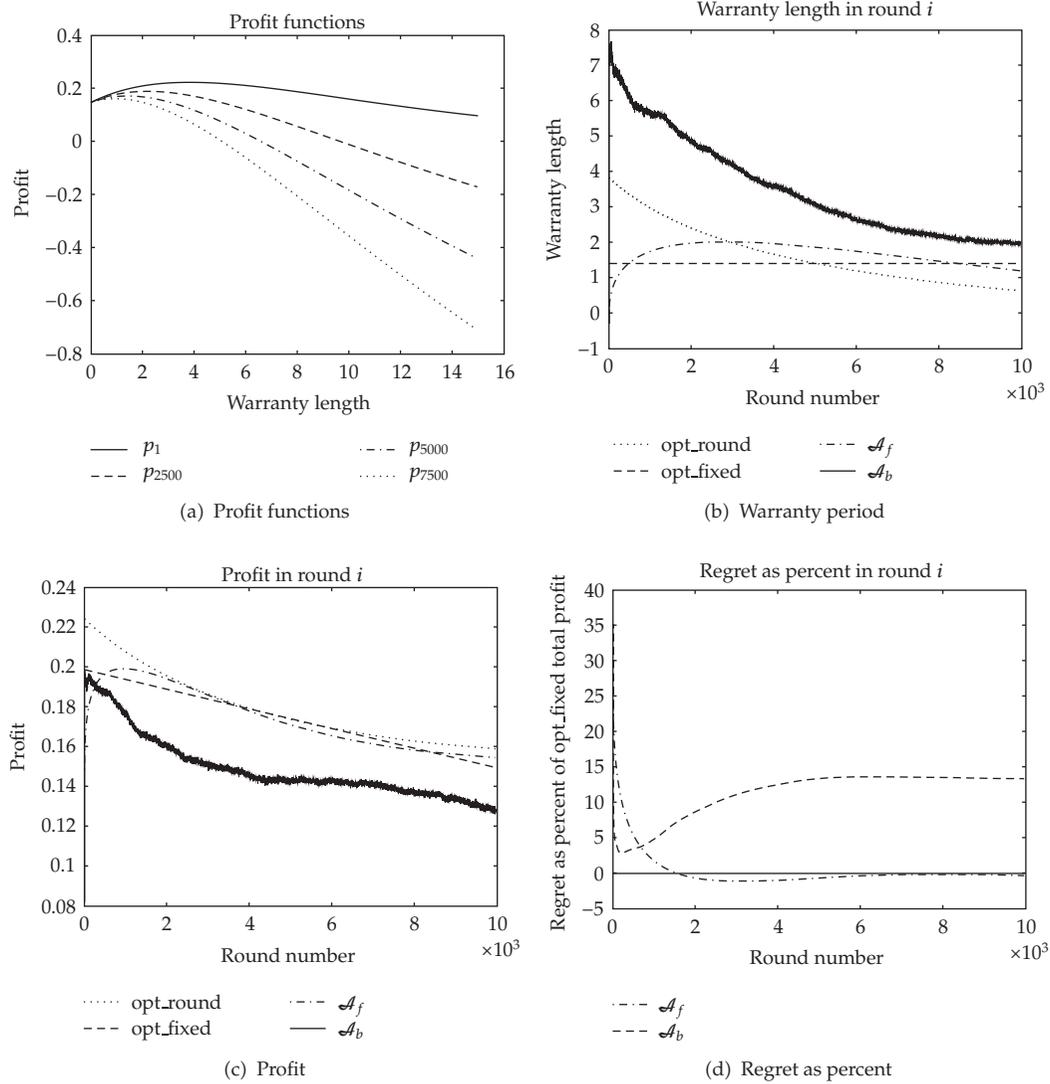(b) Warranty period

(c) Profit

(d) Regret as percent

Figure 5: Additive increase in penalties.

and decreases back toward zero. It can also be seen that $\mathcal{A}_b$ starts with a warranty of 7.5 and decreases toward zero. Figure 5(c) shows a similar performance of "opt_round", "opt_fixed", and $\mathcal{A}_f$. In that figure, it is clear that $\mathcal{A}_b$ starts with a poor performance, but in the long run approaches the performance of "opt_fixed". The graph in Figure 5(d) can be explained through understanding the performance of algorithm $\mathcal{A}_b$, which is outlined in the appendices.

In our penalty example, the optimal warranty approaches zero quickly. So, algorithm $\mathcal{A}_b$ cannot set a warranty close to zero because of the algorithm's projection to a subset of the feasible region. As the algorithm's parameter α approaches zero, $\mathcal{A}_b$ can set a warranty closer and closer to zero. Thus, we can expect the regret shown in Figure 5(d) to decrease toward zero at a speed of $O(\sqrt[6]{n})$, matching that of the parameter α.

5. Conclusions and Future Research Directions

In this paper we have presented a framework for analysis of warranty using an online convex optimization algorithm. We have introduced a class of profit functions that can be used to model a competitive market with warranties. We have shown that under incomplete information regarding the future changes in the environment, the decision maker could choose a warranty strategy that achieves a profit similar to the profit, that could have been generated by the unknown optimal warranty. In specific, we use the results of Zinkevich and Flaxman et al. to exhibit strategies achieving near optimal profits, that is, strategies with regret approaching zero in a long term. We exhibit several settings of changing environment and show that in each of these, the online algorithms can provide a reasonable support in warranty-related decision making.

This study demonstrates that it is feasible for a company to maximize profit through adjusting warranty in a dynamic environment, without knowledge of the current or future market conditions. However, the algorithms presented here do have explicit limitations that should be noted before use in a real environment. First, as most optimization algorithms, the algorithms presented in the paper are guaranteed to work for convex objective functions. However, if the profit function of the company is not convex, it is possible for the algorithm to get stuck in a local optimum. Furthermore, as mentioned earlier, some products, such as cars, may not be appropriate for use with these algorithms because of the real-time length of a round, which is on the order of years. As demonstrated, specifically for the bandit algorithm, a large number of rounds are required to approach the optimal warranty period.

Furthermore, we are able to identify two possible directions for further research. One option is to focus on reducing the limitations of the used online algorithms. It would be interesting to see if these algorithms can be coupled with existing algorithms for avoiding local optima. For example, is it possible to pair the bandit algorithm with simulated annealing? What would such a pairing do to the regret guarantees of the original bandit algorithm? Would such a pairing deliver good performance in avoiding local optima? Another possible direction for further research is to try to apply our results to a real data; related to the performance of brokerage firms. Firstly, it will be challenging to find the appropriate set of real data. Moreover, it would be interesting to come up with a method for estimating the parameters of the profit function from real data; parameters such as the total market size, the failure CDF, and the market share as a function of warranty period. Such an estimation would make it possible to investigate the application of these algorithms in a realistic situation.

Appendices

A. The Online Algorithm

As mentioned earlier, we concern ourselves with profit maximization. Thus, consider an online convex programming problem consisting of a maximization objective, a feasible region $F \in \Re^n$, and an infinite sequence of concave functions $\{p_1, p_2, \ldots, p_i, \ldots\}$, each going from F to $\Re$. We present the main ideas of the online algorithm in two different settings: firstly, in full-information settings, when the profit function p_i is fully known after each round and secondly, in bandit settings, when the profit function p_i is unknown, and only its value $p_i(w_i)$ is revealed after the ith round.

Assumptions and Definitions

(i) The feasible region $F \in \mathfrak{R}^n$ is

(1) a bounded set, that is, for any $x, y \in F$, there exists $N \in \mathfrak{R}$, so that $d(x, y) \leq N$, where $d(x, y) = \|x - y\|$ and $\|x\| = \sqrt{x \cdot x}$;

(2) a closed set, that is, for any sequence $\{w_i\}_1^\infty$, $w_i \in F$, if there exists $x \in \mathfrak{R}^n$ such that $x = \lim_{i \to \infty} w_i$, then $x \in F$;

(3) a nonempty set;

(4) a convex set.

(ii) The profit functions are differentiable.

(iii) There exists $N \in \mathfrak{R}$, so that for all i and for all $x \in F$, $\|\nabla p_i(x)\| \leq N$.

(iv) For all $y \in \mathfrak{R}^n$, there exists an algorithm to produce $\arg\min_{x \in F} d(x, y)$.

(v) For all i, there exists an algorithm, given x, to get $\nabla p_i(x)$.

(vi) The projection of y over F is $P(y) = \arg\min_{x \in F} d(x, y)$.

(vii) The regret of $\mathcal{A}$ until T is $R_{\mathcal{A}}(T) = (\max_{x^* \in F} \sum_{i=1}^T p_i(x^*)) - \sum_{i=1}^T p_i(w_i)$.

(viii) A function $p(x)$ satisfies an L-Lipschitz condition if there exists a real constant L such that $d(p(x), p(y)) \leq L d(x, y)$.

B. Online Gradient Descent Algorithm $\mathcal{A}_f$ in Full-Information Settings

Assume that the profit function p_i is fully known after the ith round. Select an initial $w_1 \in F$ and an updating sequence $\eta = \{\eta_1, \eta_2, \ldots, \eta_i, \ldots\}$ with each $\eta_i \in \mathfrak{R}^+$. In time step $(i + 1)$, after evaluating the profit function $p_i(w_i)$, move to the next point, which is

$$\mathcal{A}_f : x_{i+1} = P(x_i - \eta_i \nabla p_i(w_i)). \tag{B.1}$$

Assuming that the updating sequence η has the form $\eta_i = 1/\sqrt{i}$, Zinkevich [5] has shown that the regret of the algorithm $\mathcal{A}_f$, given in (B.1), is

$$R_{\mathcal{A}_f}(T) \leq \frac{\|F\|^2 \sqrt{T}}{2} + \left(\sqrt{T} - \frac{1}{2}\right)\|\nabla p\|^2. \tag{B.2}$$

Therefore,

$$\limsup_{T \to \infty} \frac{R_{A_f}(T)}{T} \leq 0, \tag{B.3}$$

where

$$\|F\| = \max_{x,y \in F} d(x, y), \qquad \|\nabla p\| = \sup_{x \in F, i \in \{1,2,\ldots\}} \|\nabla p_i(x)\|^2. \tag{B.4}$$

Imposing stronger assumptions on the profit functions and choosing appropriately the step sizes, Hasan et al. [8] have extended Zinkevich's ideas by proposing several algorithms achieving logarithmic regret.
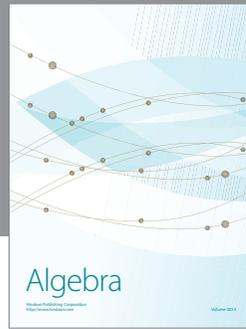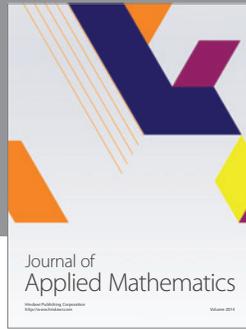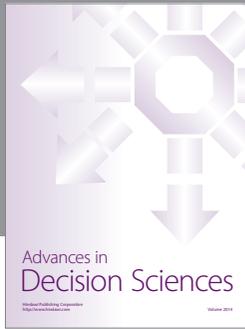
C. Online Gradient Descent Algorithm $\mathcal{A}_b$ in Bandit Settings

In bandit settings, after the ith round, the profit function p_i is unknown, and only its value $p_i(w_i)$ is revealed. Therefore, the gradient of p_i, needed for $\mathcal{A}_f$, cannot be accessed directly. The main difficulties in bandit setting is to obtain a one-point estimate of the gradient $\nabla p_i(w_i)$. Algorithm $\mathcal{A}_b$ works as follows. It has a sequence of points y_i at which it would like to perform gradient descent, as in algorithm $\mathcal{A}_f$. However, to estimate the gradient at y_i, $\mathcal{A}_b$ select a uniformly random point w_i from a small circle around y_i. Algorithm $\mathcal{A}_b$ then sets y_{i+1} to y_i shifted in the direction of w_i with distance proportional to $p_i(w_i)$. To be sure that y_{i+1} is in the feasible region, the algorithm does a projection to a subset of the feasible region that has a small border around it. The reason for this projection to a subset is that future estimates of the gradient using randomly chosen points in a small circle should be entirely contained in the feasible region.

The algorithm then has three main parameters that change as the round number increases. Using the notation of Flaxman et al., the first parameter δ denotes the radius of the small circle around y_i from which we choose a uniformly random point w_i. The second parameter ν denotes the distance with which we move in the direction of the chosen point w_i. And the final parameter α denotes the border that we keep around the subset of the feasible region. Each of these parameters goes to zero as the round number increases. The parameters δ, ν, and α go to zero at speeds of $O(\sqrt[3]{n})$, $O(\sqrt{n})$, and $O(\sqrt[6]{n})$, respectively. Flaxman et al. [6] have shown that if the profit functions are L-Lipschitz the guarantee on the expected regret of $\mathcal{A}_b$ is $O(T^{3/4})$. Moreover, if no Lipschitz or bounded gradient assumptions are placed on the profit functions, the guarantee on the expected regret is $O(T^{5/6})$. For more details, see Flaxman et al. [6].

References

[1] W. Blischke and D. N. P. Murthy, *Warranty Cost Analysis*, Marcel Dekker, New York, NY, USA, 1993.

[2] W. Blischke and D. N. P. Murthy, *Product Warranty Handbook*, Marcel Dekker, New York, NY, USA, 1996.

[3] S. Chukova and Y. Hayakawa, "Warranty cost analysis: non-zero repair time," *Applied Stochastic Models in Business and Industry*, vol. 20, no. 1, pp. 59–71, 2004.

[4] S. Chukova and Y. Hayakawa, "Warranty cost analysis: renewing warranty with non-zero repair time," *International Journal of Reliability, Quality and Safety Engineering*, vol. 11, no. 2, pp. 1–20, 2004.

[5] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proceedings of the 20th International Conference on Machine Learning (ICML '03)*, pp. 928–936, Washington, DC, USA, August 2003.

[6] A. D. Flaxman, A. T. Kalai, and H. B. McMahan, "Online convex optimization in the bandit setting: gradient descent without a gradient," in *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '05)*, pp. 385–394, Vancouver, Canada, January 2005.

[7] D. Bell, R. Keeney, and J. Little, "Market share theorem," *Journal of Marketing Research*, vol. 12, no. 2, pp. 136–141, 1975.

[8] E. Hasan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Machine Learning*, vol. 96, no. 2-3, pp. 169–192, 2007.