Ellipsoids for Contextual Dynamic Pricing

MAXIME C. COHEN
Stern School of Business, New York University
and
ILAN LOBEL
Stern School of Business, New York University
and
RENATO PAES LEME
Google Research NYC

We study the dynamic pricing problem faced by a firm selling differentiated products. At each period, the firm receives a new product, which is described by a vector of features. The firm needs to choose prices, but it does not know a priori the market value of the different features. We first consider an algorithm that we call PolytopePricing, but prove that it incurs worst-case regret that scales exponentially in the dimensionality of the feature space. We then consider a closely related algorithm, EllipsoidPricing, and show it incurs low regret with regards to both the time horizon and the dimensionality of the feature space. For more details, we refer the reader to our full paper.

1. INTRODUCTION

In this letter, we briefly survey the results of our recent paper, [Cohen et al. 2016]. We also discuss recent developments in the literature surrounding our paper.

We consider an online market where a firm sells highly differentiated products to its buyers. In each period, a new product arrives and the selling firm must set a price for it. Each product is characterized by a vector of features (or contexts) that determine its market value. The firm knows the features of the products, but it does not know a priori the value of the different features. Our paper aims to understand what is a good pricing policy for balancing learning and earning in such a contextual setting.

Our problem is motivated by online markets such as Airbnb and the market for online ads. In these markets, every product is unique in its attributes. In the case of Airbnb, a product is a stay in a particular listing on a specific date. The features therefore represent both characteristics of the listing, such as location and amenities, as well as the check-in/check-out dates. In the market for online ads, a product is an impression which is sold to potential advertisers. The features that determine the market value of an impression include the IP address and the relevant cookie data, which might contain information such as gender, age and browsing history.

Authors' addresses: maxime.cohen@stern.nyu.edu, ilobel@stern.nyu.edu, renatoppl@google.com

2. THE MODEL

Consider a seller that receives products in an online fashion. At every period $t = 1, ..., T$, a new product arrives with a set of features $x_t \in \mathcal{X} \subset \mathbb{R}^d$ such that $||x_t||_2 \leq 1$. The market value of each feature is given by $v_t = \theta' x_t$, where θ is a d-dimensional vector unknown to the seller. The seller chooses a price p_t as a function of x_t. If the seller chooses a price below or equal to v_t, a sale occurs and she earns p_t. If the seller chooses a price p_t above v_t, no transaction occurs.

The seller knows initially only that θ belongs to a bounded convex set K_1, where $||\theta||_2 \leq 1$. If a sale occurs at time t, the seller updates her uncertainty set according to $K_{t+1} = K_t \cap \{\theta \in \mathbb{R}^d : \ \theta' x_t \geq p_t\}$. Similarly, if a sale does not occur at time t, the seller updates her uncertainty set following $K_{t+1} = K_t \cap \{\theta \in \mathbb{R}^d : \ \theta' x_t \leq p_t\}$. The seller's problem is to choose a pricing policy that minimizes her worst-case regret, which is given by:

$$\text{REGRET} = \max_{\theta \in K_1, \ \{x_t\}} \sum_{t=1}^{T} \Big[\theta' x_t - p_t I\{\theta' x_t \geq p_t\}\Big]. \tag{1}$$

Our model assumes that both θ and the feature vectors $\{x_t\}$ are chosen adversarially by nature. Good policies under this metric would therefore be robust to changes or seasonal fluctuation in the pattern of arrival of the feature vectors over time.

The model as described above assumes a linear deterministic relationship between the features and the market values. We also briefly discuss below how to address more general market value functions, including both noisy valuation models and some commonly used nonlinear models.

3. CONTEXTUAL PRICING ALGORITHMS

If the seller knew the value of θ, she could maximize her revenue by simply choosing $p_t = \theta' x_t$ at each period. However, since the seller does not know the value of θ, she must balance exploration and exploitation. Since the optimal decision at each period depends on a context vector x_t, our problem is a special case of a contextual bandit problem ([Auer 2003]). We could therefore use an off-the-shelf algorithm for contextual bandits, such as [Agarwal et al. 2014]. Such an algorithm would have suboptimal performance (polynomial rather than logarithmic regret in T) since it does not take advantage of the underlying linear structure of our pricing problem. Consequently, our aim is to construct an algorithm for this problem with good performance with respect to both the time horizon T and the dimensionality d. Our first attempt is a multidimensional version of binary search that we call PolytopePricing.

The PolytopePricing Algorithm. At each period t, we have access to the vector of features x_t and we know that $\theta \in K_t$. A natural first question is to ask whether we can predict the value $v_t = \theta' x_t$ with reasonable accuracy. The lowest and highest possible values of v_t are given by:

$$\underline{b}_t = \min_{\hat{\theta} \in K_t} \hat{\theta}' x_t \qquad \text{and} \qquad \bar{b}_t = \max_{\hat{\theta} \in K_t} \hat{\theta}' x_t.$$

For a given accuracy parameter $\epsilon > 0$, we can say that we know the value of v_t with ϵ-accuracy if and only if $\bar{b}_t - \underline{b}_t \leq \epsilon$.

This gives rise to a parameterized algorithm PolytopePricing(ϵ). If $\bar{b}_t - \underline{b}_t \le \epsilon$, then we should choose an exploit price since we know the market value of product t with ϵ-accuracy. The natural choice for an explore is $p_t = \underline{b}_t$ since no price above this level guarantees a sale. If $\bar{b}_t - \underline{b}_t > \epsilon$, then we should use an explore price. A natural explore price is the one inspired by binary search: $p_t = (\bar{b}_t + \underline{b}_t)/2$.

Unfortunately, as shown in Theorem 1 below, this algorithm does not scale well with the dimensionality of the feature space. This algorithm becomes problematic in high dimensions because the explore price might remove too small a fraction of the uncertainty set.

Theorem 1. *For any parameter $\epsilon > 0$, the algorithm* PolytopePricing *suffers worst-case regret* $\Omega(1.2^d)$.

The EllispoidPricing Algorithm. A natural follow-up question is whether we can "fix" PolytopePricing by ensuring that we remove a sufficiently large volume of the uncertainty set per explore period. A way to accomplish this objective is to borrow ideas from the ellipsoid method from optimization theory ([Khachiyan 1979]).

The EllispoidPricing algorithm works exactly as the PolytopePricing, except it has an additional step. At the end of each period t, we replace our convex set K_t by its Löwner-John ellipsoid E_t. The Löwner-John ellipsoid of a convex set is the smallest ellipsoid that contains that set. It turns out that this simple modification to the algorithm is sufficient to ensure a good regret performance.

Theorem 2. *The worst-case regret of the* EllipsoidPricing *algorithm with parameter $\epsilon = d^2/T$ is $O(d^2 \ln(T/d))$.*

To prove Theorem 2, we combine two ideas. From the work of Khachiyan, we know that an algorithm that iteratively cuts an ellipsoid and then replaces the remaining half-ellipsoid by its own Löwner-John ellipsoid yields an exponentially fast volume reduction. This idea alone is not sufficient to prove our result since our theorem requires us to control the radii of the ellipsoid too, not merely the volume. Note that bounding the radii of the period t ellipsoid allows us to bound the difference $\bar{b}_t - \underline{b}_t$, and thus determines the accuracy of our knowledge of the value of v_t. This now brings us to the second part of the proof. Building on the linear algebra machinery of [Wilkinson 1965], we prove that the radii of our ellipsoids never become too small. This follows from the fact that our algorithm never uses an explore price in a direction that is already small; it uses an exploit price instead.

The EllipsoidPricing algorithm not only has a good worst-case regret guarantee, it is also computationally efficient. Both key operations — optimizing a linear function over an ellipsoid and finding the Löwner-John ellipsoid of a half-ellipsoid — require only matrix-vector multiplications.

Nonlinear and noisy market values. The model above assumes the market value is a deterministic linear function of the feature vector. Our algorithm can be extended to more general models. If the market value function can be expressed as $v_t = f(\theta' \phi(x_t))$, where f is a non-decreasing Lipschitz continuous function and $||\phi(\cdot)||_2 \le 1$, then Theorem 2 still applies up to the Lipschitz constant.

If the market value is a noisy function of the feature vector, i.e., $v_t = \theta' x_t + \delta_t$ for a noise term δ_t, then we need to slightly modify the EllipsoidPricing algorithm before applying it. The key idea here is to add a safety margin when updating the uncertainty set. Instead of cutting the ellipsoid through its center, we can cut the ellipsoid in a way that leaves more than half of the original volume. This is called a shallow cut. In [Cohen et al. 2016], we show the details of how shallow cuts can be used to handle both bounded adversarial noise and i.i.d. Gaussian noise.

4. RECENT DEVELOPMENTS IN THE LITERATURE

There is a large body of literature on dynamically adjusting prices both with and without contextual information. We refer to [Cohen et al. 2016] for an extensive discussion. Here, we focus on papers that have appeared in the last few months and modify the model of [Cohen et al. 2016] in interesting new directions.

Lower bound. One natural question is whether our EllipsoidPricing algorithm has optimal regret. The answer is no. [Kleinberg and Leighton 2003] showed that in a one-dimensional (non-contextual) version of our problem, the optimal regret is $\Theta(\ln \ln T)$. The best known lower bound for our problem is thus $\Omega(d \ln \ln T)$.

In recent work, [Lobel et al. 2016] make progress towards closing this gap by showing that there exists an algorithm that incurs $O(d \ln(dT))$ worst-case regret. This algorithm also runs in polynomial time, but is far more complex and harder to implement than EllipsoidPricing. It requires computing approximate centroids of high-dimensional sets as well as projecting and cylindrifying polytopes.

Stochastic versions. Another direction is to consider a stochastic version of our model, as opposed to our adversarial framework. This approach harks back to [Amin et al. 2014], who proposed using stochastic gradient descent. Recently, different authors have taken this problem in very different directions. [Javanmard and Nazerzadeh 2016] show that a regularized maximum likelihood algorithm can be used to solve this problem and obtain strong performance bounds as a function of the sparsity of the feature space. In a different direction, [Qiang and Bayati 2016] show that a method as simple as greedy least squares regression performs well since the contexts (or covariates in their paper) ensure the seller performs sufficient exploration.

REFERENCES

Agarwal, A., Hsu, D., Kale, S., Langford, J., Li, L., and Schapire, R. E. 2014. Taming the monster: A fast and simple algorithm for contextual bandits. In *Proceedings of ICML*.

Amin, K., Rostamizadeh, A., and Syed, U. 2014. Repeated contextual auctions with strategic buyers. In *Proceedings of NIPS*. 622–630.

Auer, P. 2003. Using confidence bounds for exploitation-exploration trade-offs. *The Journal of Machine Learning Research 3*, 397–422.

Cohen, M., Lobel, I., and Leme, R. P. 2016. Feature-based dynamic pricing. Working paper, New York University.

Javanmard, A. and Nazerzadeh, H. 2016. Dynamic pricing in high-dimensions. Working paper, University of Southern California.

Khachiyan, L. G. 1979. A polynomial algorithm in linear programming. *Doklady Akademii Nauk SSSR 244*, 1093–1096.

Kleinberg, R. and Leighton, T. 2003. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *Proceedings of FOCS*. 594–605.

LOBEL, I., LEME, R. P., AND VLADU, A. 2016. Multidimensional binary search for contextual decision-making. Working paper, New York University.

QIANG, S. AND BAYATI, M. 2016. Dynamic pricing with demand covariates. Working Paper, Stanford University.

WILKINSON, J. H. 1965. *The Algebraic Eigenvalue Problem.* Vol. 87. Clarendon Press Oxford.