Low-Rank Bandit Methods for High-Dimensional Dynamic Pricing

Jonas Mueller, Vasilis Syrgkanis, Matt Taddy

Amazon Microsoft Research Amazon

jonasmueller@csail.mit.edu

Summary

- Dynamic pricing $\approx$ online bandit optimization
- Bandit pricing can handle changing (adversarial) environments
- Number of pricing rounds until standard bandits find optimal prices (vanishing regret) depends on # of products
- Observed product demands provide side information if they evolve in a low-rank fashion based on (latent) product features
- We develop bandit pricing algorithms that exploit this assumption, whose regret vanishes at a rate that only depends on # of product features instead of # of products

Objective

Choose prices for many products to maximize revenue/profit. Update prices over time to reflect changing demand curves.

Setup

- $\mathbf{p}_t$ = vector of prices for N products sold during time-period t
- $\mathbf{q}_t \in \mathbb{R}^N$ = vector of demands for each product in time-period t
- $R_t(\mathbf{p}_t)$ = total revenue over period $t = \langle \mathbf{q}_t, \mathbf{p}_t \rangle$
- Regret $= \mathbb{E}\left[\sum_{t=1}^{T} R_t(\mathbf{p}^*) - R_t(\mathbf{p}_t)\right]$
- $\mathbf{p}^*$ = optimal price vector (chosen in hindsight)

Standard Demand Model

$$\mathbf{q}_t = \mathbf{c}_t - \mathbf{B}_t \mathbf{p}_t + \boldsymbol{\epsilon}_t$$

- $\mathbf{B}_t$ describes how price of products affects demand for other products in round t (asymmetric, positive definite matrix)

- Can achieve $O(T^{3/4} N^{1/2})$ regret using standard online bandit method to select prices under this model
- Flaxman, Kalai, McMahan (2005): *Online convex optimization in the bandit setting: Gradient descent without a gradient*
- **Idea**: Set price $\mathbf{p} + \delta\boldsymbol{\xi}$ instead of $\mathbf{p}$ with $\boldsymbol{\xi}$ = random unit vector. For $\tilde{R}(\mathbf{p}) := R(\mathbf{p} + \delta\boldsymbol{\xi})$: $\nabla \tilde{R}(\mathbf{p}) = \mathbb{E}[R(\mathbf{p} + \delta\boldsymbol{\xi})\boldsymbol{\xi}]N/\delta$

Product Features

- Let $\mathbf{u}_i$ = d-dimensional features of product i $\quad (d \ll N)$
- Product similarity = $\langle \mathbf{u}_i, \mathbf{u}_j \rangle_{\mathbf{V}_t}$, effect of p_j on $q_i = \mathbf{u}_i^T \mathbf{V}_t \mathbf{u}_j \cdot p_j$

Low-Rank Demand Model

$$\mathbf{q}_t = \mathbf{U}\mathbf{z}_t - \mathbf{U}\mathbf{V}_t\mathbf{U}^T\mathbf{p}_t + \boldsymbol{\epsilon}_t$$

- $\mathbf{U} = N \times d$ matrix, whose rows = product features

Known Product Features

- Algebra $\implies$ $R_t(\mathbf{p}) = f_t(\mathbf{x})$ for convex f_t and $\mathbf{x} := \mathbf{U}^T\mathbf{p} \in \mathbb{R}^d$
- **Strategy:** use bandit algorithm to optimize $\mathbf{x}_t$ w.r.t. f_t, each time selecting prices $\mathbf{p}_t$ via pseudo-inverse s.t. $\mathbf{x}_t = \mathbf{U}^T\mathbf{p}_t$
- Regret $= O(T^{3/4} d^{1/2})$ (does not depend on N)

Unknown Product Features

- Assume orthonormal product features: $\mathbf{U}\mathbf{x} = \mathbf{p}$ for $\mathbf{x} := \mathbf{U}^T\mathbf{p}$
- Previous strategy does not need known $\mathbf{U}$, only need $\text{span}(\mathbf{U})$
- If $\boldsymbol{\epsilon}_t = 0$, then $\text{span}(\mathbf{U}) = \text{span}(\mathbf{q}_1, \dots, \mathbf{q}_d)$
- When $\boldsymbol{\epsilon}_t \neq 0$, we can estimate $\text{span}(\mathbf{U})$ via SVD of $[\mathbf{q}_1, \dots, \mathbf{q}_t]$
- Run bandit algorithm to optimize $\mathbf{x}_t = \widehat{\mathbf{U}}^T\mathbf{p}_t$ w.r.t. f_t where $\text{span}(\widehat{\mathbf{U}})$ = current estimate of $\text{span}(\mathbf{U})$
- Regret $= O(T^{3/4} d)$ (does not depend on N)

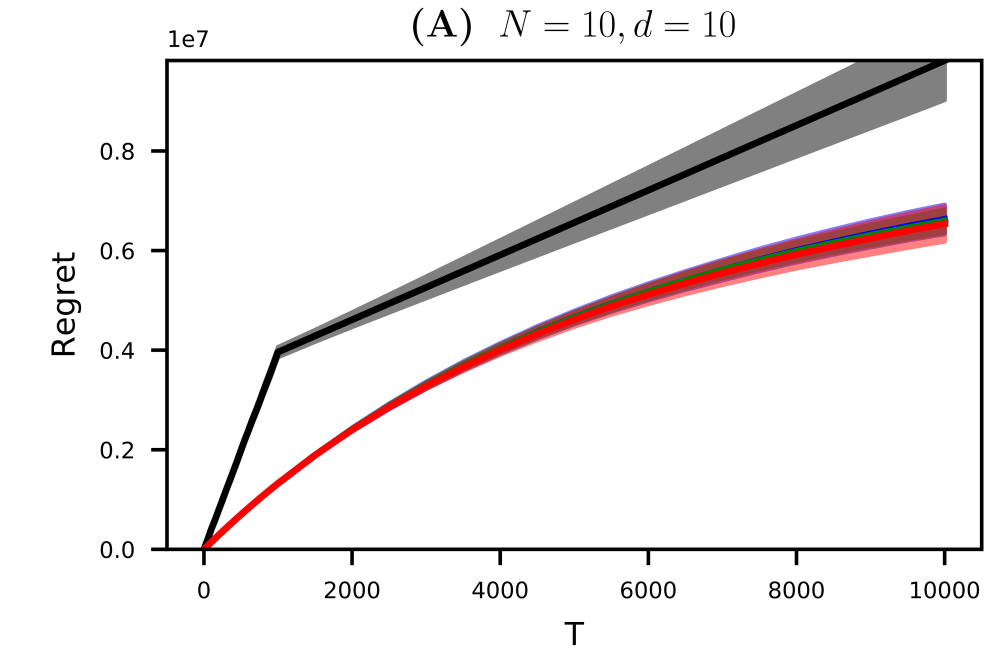
Algorithm 1 Online Pricing Optimization with Latent Features

Input: $\eta, \delta, \alpha > 0$, rank $d \in [1, N]$, initial prices $\mathbf{p}_0 \in \mathcal{S}$
Output: Prices $\mathbf{p}_1, \dots, \mathbf{p}_T$ to maximize overall revenue

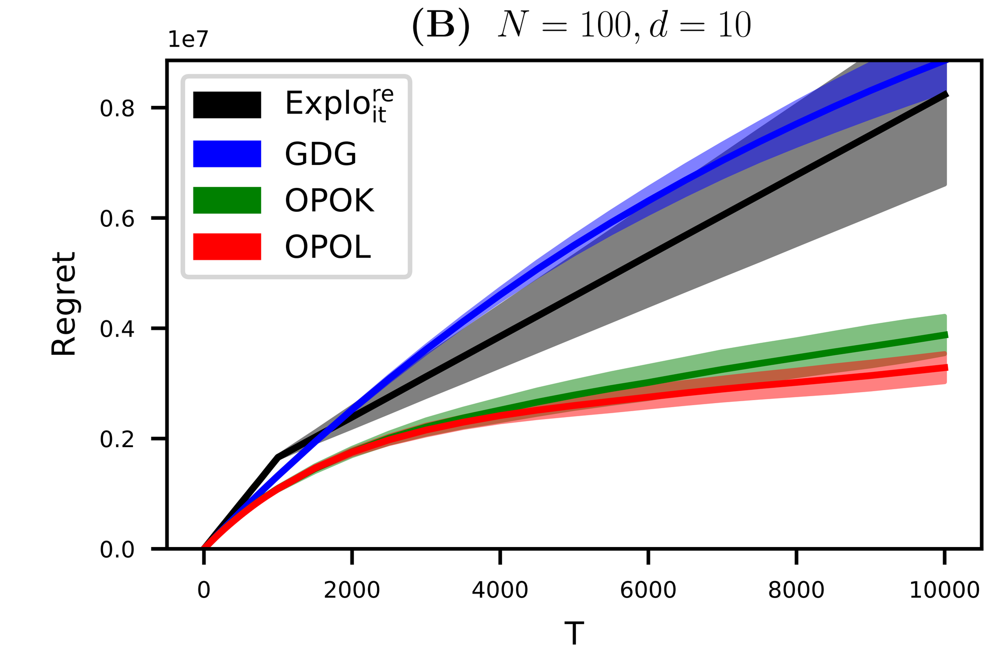
Initialize $\widehat{\mathbf{Q}}$ as $N \times d$ matrix of zeros
Initialize $\widehat{\mathbf{U}}$ as random $N \times d$ orthogonal matrix
Set prices to $\mathbf{p}_0 \in \mathcal{S}$ and observe $\mathbf{q}_0(\mathbf{p}_0), R_0(\mathbf{p}_0)$
Define $\mathbf{x}_1 = \widehat{\mathbf{U}}^T\mathbf{p}_0$
for $t = 1, \dots, T$:
 $\tilde{\mathbf{x}}_t := \mathbf{x}_t + \delta\boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t \sim \text{Unif}(\{\mathbf{x} \in \mathbb{R}^d : ||\mathbf{x}||_2 = 1\})$
 Set prices: $\mathbf{p}_t = \widehat{\mathbf{U}}\tilde{\mathbf{x}}_t$ and observe $\mathbf{q}_t(\mathbf{p}_t), R_t(\mathbf{p}_t)$
 $\mathbf{x}_{t+1} = \text{PROJECTION}\big(\mathbf{x}_t - \eta R_t(\mathbf{p}_t)\boldsymbol{\xi}_t \rightarrow (1-\alpha)\cdot\widehat{\mathbf{U}}^T\mathcal{S}\big)$
 $\widehat{\mathbf{Q}}_{*,j} \leftarrow \frac{1}{k}\mathbf{q}_t + \frac{k-1}{k}\widehat{\mathbf{Q}}_{*,j}$ with $j = 1 + [(t-1) \bmod d]$, $k = \lfloor t/d \rfloor$
 Set columns of $\widehat{\mathbf{U}}$ as top d left singular vectors of $\widehat{\mathbf{Q}}$

Stationary Demand Model

(A) $N = 10, d = 10$

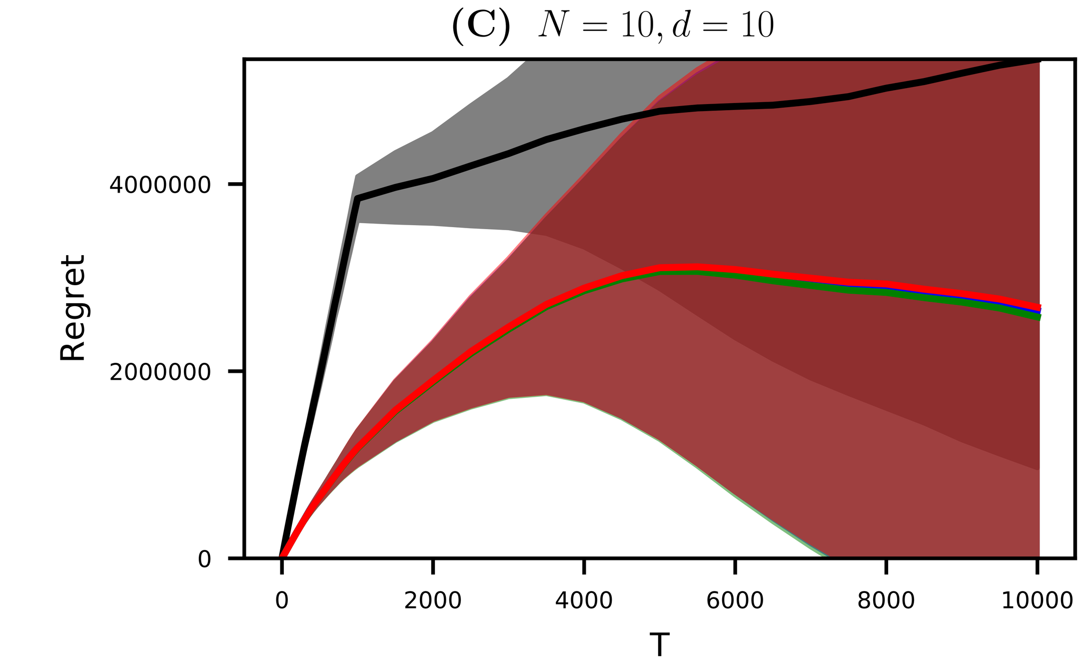


(B) $N = 100, d = 10$



Evolving Demand Model

(C) $N = 10, d = 10$



(D) $N = 100, d = 10$