Smoothness-Adaptive Dynamic Pricing with Nonparametric Demand Learning

Hansheng Jiang
Rotman School of Management
University of Toronto

Joint work with Zeqi Ye (Nankai→Northwestern)

INFORMS RMP Conference
Los Angeles

Dynamic Pricing: Learning While Earning

One of the core problems in modern revenue management

Dynamic Pricing: Learning While Earning

One of the core problems in modern revenue management

Examples: online retailing, transportation (ride-hailing, airline), hospitality (hotel pricing)

Dynamic Pricing: Learning While Earning

One of the core problems in modern revenue management

Examples: online retailing, transportation (ride-hailing, airline), hospitality (hotel pricing)

Typical procedure in practice

1. Collect price and demand data
2. Estimate a parametric or nonparametric demand model
3. Optimize price based on current demand estimates
4. Observe consumer responses to prices and repeat the above two steps

Dynamic Pricing: Learning While Earning

One of the core problems in modern revenue management

Examples: online retailing, transportation (ride-hailing, airline), hospitality (hotel pricing)

Typical procedure in practice

1. Collect price and demand data
2. Estimate a parametric or nonparametric demand model
3. Optimize price based on current demand estimates
4. Observe consumer responses to prices and repeat the above two steps

Naturally, the estimation method and pricing performances correspond to demand assumptions

Model

- **Price-setting monopolist** decision variable $p_t \in \mathcal{P}$

Model

- **Price-setting monopolist** decision variable $p_t \in \mathcal{P}$
- **Consumer demand** random demand d_t and $\mathbb{E}[d_t | p_t = p] = f(p)$

Model

- **Price-setting monopolist** decision variable $p_t \in \mathcal{P}$

- **Consumer demand** random demand d_t and $\mathbb{E}[d_t | p_t = p] = f(p)$
  - nonparametric demand with regularity assumptions

Model

- **Price-setting monopolist** decision variable $p_t \in \mathcal{P}$
- **Consumer demand** random demand d_t and $\mathbb{E}[d_t | p_t = p] = f(p)$
  - nonparametric demand with regularity assumptions
- **Pricing policy** $\pi = (p_1, p_2, \dots)$ non-anticipating

Model

- **Price-setting monopolist** decision variable $p_t \in \mathcal{P}$

- **Consumer demand** random demand d_t and $\mathbb{E}[d_t | p_t = p] = f(p)$
  - nonparametric demand with regularity assumptions

- **Pricing policy** $\pi = (p_1, p_2, \dots)$ non-anticipating

- **Regret** $R^\pi(T) = \mathbb{E}^\pi \left[\sum_{t=1}^T \{p^* f(p^*) - p_t f(p_t)\} \right]$

Previous Results

Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters

Previous Results

Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters
- Revenue function $p \cdot f(p)$ is concave and thus unimodal

Previous Results

Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters
- Revenue function $p \cdot f(p)$ is concave and thus unimodal
- Optimal regret rate is $\widetilde{O}(\sqrt{T})$

Previous Results

Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters
- Revenue function $p \cdot f(p)$ is concave and thus unimodal
- Optimal regret rate is $\widetilde{O}(\sqrt{T})$

Lipschitz demand model

- Demand function $f(p)$ is Lipschitz

Previous Results

Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters
- Revenue function $p \cdot f(p)$ is concave and thus unimodal
- Optimal regret rate is $\widetilde{O}(\sqrt{T})$

Lipschitz demand model

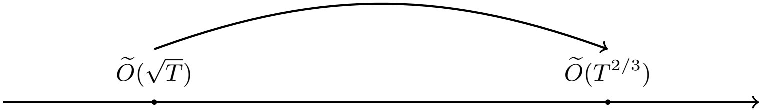
- Demand function $f(p)$ is Lipschitz
- Optimal regret rate is $\widetilde{O}(T^{2/3})$

Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters
- Revenue function $p \cdot f(p)$ is concave and thus unimodal
- Optimal regret rate is $\widetilde{O}(\sqrt{T})$

Lipschitz demand model

- Demand function $f(p)$ is Lipschitz
- Optimal regret rate is $\widetilde{O}(T^{2/3})$

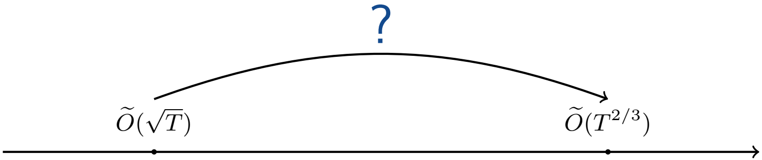


Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters
- Revenue function $p \cdot f(p)$ is concave and thus unimodal
- Optimal regret rate is $\widetilde{O}(\sqrt{T})$

Lipschitz demand model

- Demand function $f(p)$ is Lipschitz
- Optimal regret rate is $\widetilde{O}(T^{2/3})$



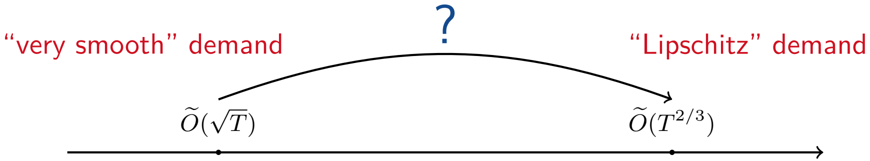
Previous Results

Canonic parametric demand model

- Linear demand $f(p) = a + b \cdot p$, where a, b are unknown parameters
- Revenue function $p \cdot f(p)$ is concave and thus unimodal
- Optimal regret rate is $\widetilde{O}(\sqrt{T})$

Lipschitz demand model

- Demand function $f(p)$ is Lipschitz
- Optimal regret rate is $\widetilde{O}(T^{2/3})$



Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot \left| p - p' \right|^{\beta - w(\beta)}$$

Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot |p - p'|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β

Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot \left| p - p' \right|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!

Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot |p - p'|^{\beta - w(\beta)}$$

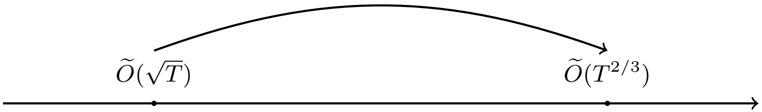
- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!

Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \le L \cdot |p - p'|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!

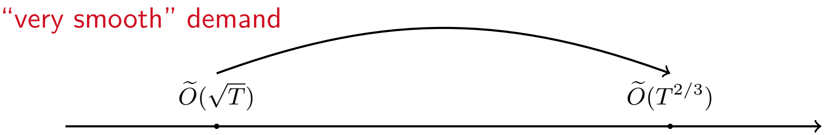


Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot |p - p'|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!

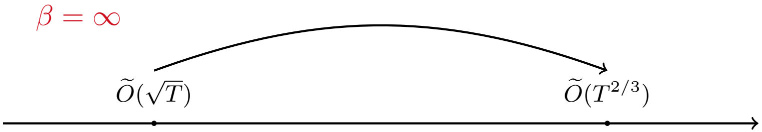


Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot |p - p'|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!

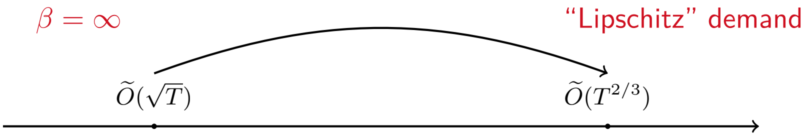


Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \le L \cdot \left| p - p' \right|^{\beta - w(\beta)}$$

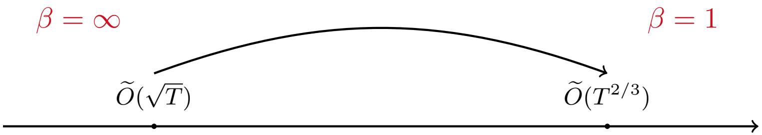
- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!



- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot |p - p'|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!

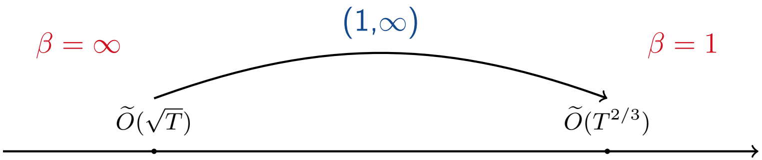


Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot |p - p'|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!

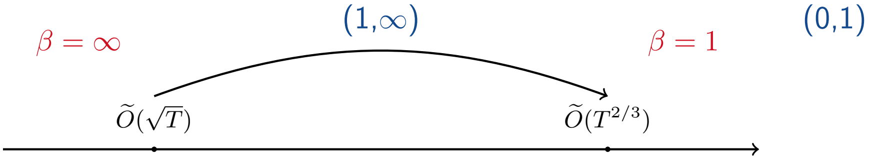


Regret Under Nonparametric Demand

- **Hölder-smooth demand function** $f(p) \in \mathcal{H}(\beta, L)$ for some $\beta, L > 0$ if

$$\left| f^{(w(\beta))}(p) - f^{(w(\beta))}(p') \right| \leq L \cdot |p - p'|^{\beta - w(\beta)}$$

- **Role of smoothness parameter** β
  - optimal regret rate $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$ rate heavily **relies** on β!
  - the algorithm design also **relies** on the knowledge of β!



Can we design a dynamic pricing strategy that **does not** require the prior knowledge of β while maintaining the **optimal** regret of $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$?

Can we design a dynamic pricing strategy that **does not** require the prior knowledge of β while maintaining the **optimal** regret of $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$?

Rely on Smoothness Parameter

Can we design a dynamic pricing strategy that **does not** require the prior knowledge of β while maintaining the **optimal** regret of $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$?

Rely on Smoothness Parameter

Can we design a dynamic pricing strategy that **does not** require the prior knowledge of β while maintaining the **optimal** regret of $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$?

Rely on Smoothness Parameter

Smoothness-Adaptive

Can we design a dynamic pricing strategy that **does not** require the prior knowledge of β while maintaining the **optimal** regret of $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$?

Rely on Smoothness Parameter

Smoothness-Adaptive

Our Contributions

- **Characterizing adaptivity challenge**
  Without additional conditions, achieving the optimal regret for functions without knowing the Hölder smoothness parameter is impossible.

Our Contributions

- **Characterizing adaptivity challenge**
  Without additional conditions, achieving the optimal regret for functions without knowing the Hölder smoothness parameter is impossible.

- **Proposing a self-similarity condition to enable adaptivity**
  This condition does not decrease problem complexity with the same regret lower bound $\Omega(T^{\frac{\beta+1}{2\beta+1}})$

Our Contributions

- **Characterizing adaptivity challenge**
  Without additional conditions, achieving the optimal regret for functions without knowing the Hölder smoothness parameter is impossible.

- **Proposing a self-similarity condition to enable adaptivity**
  This condition does not decrease problem complexity with the same regret lower bound $\Omega(T^{\frac{\beta+1}{2\beta+1}})$

- **Optimal minimax regret rate**
  Our Smoothness-Adaptive Dynamic Pricing (`SADP`) algorithm enjoys an optimal regret bound $\widetilde{O}(T^{\frac{\beta+1}{2\beta+1}})$

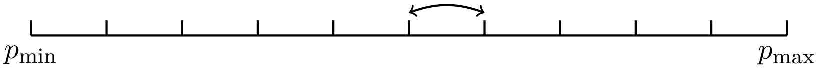
Role of Smoothness β

- If smoothness parameter β **is known**, optimal regret rate is achievable by our Hölder-Smooth Dynamic Pricing (HSDP) Algorithm

Role of Smoothness β

- If smoothness parameter β **is known**, optimal regret rate is achievable by our Hölder-Smooth Dynamic Pricing (`HSDP`) Algorithm

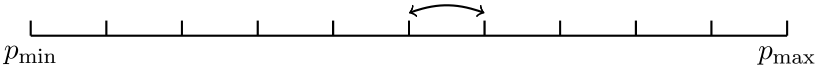
`HSDP` Algorithm Design



Role of Smoothness β

- If smoothness parameter β **is known**, optimal regret rate is achievable by our Hölder-Smooth Dynamic Pricing (`HSDP`) Algorithm

`HSDP` Algorithm Design



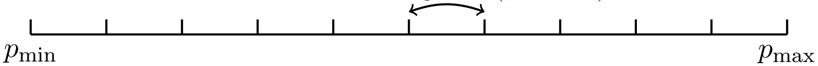
Local polynomial regression $\downarrow$ (in each small interval) + Multi-armed bandits $\downarrow$ (select the "optimal" interval)

Role of Smoothness β

- If smoothness parameter β **is known**, optimal regret rate is achievable by our Hölder-Smooth Dynamic Pricing (`HSDP`) Algorithm

`HSDP` **Algorithm Design**

interval length $O(T^{-\frac{1}{2\beta+1}})$



Local polynomial regression $\downarrow$ (in each small interval)

+

Multi-armed bandits $\downarrow$ (select the "optimal" interval)

Characterizing Adaptivity Challenge

- It is impossible to achieve adaption without additional assumptions

Characterizing Adaptivity Challenge

- It is impossible to achieve adaption without additional assumptions

Theorem

Fix any two positive Hölder smoothness parameters $\alpha > \beta > 0$*, and parameters* $L(\alpha), L(\beta) > 0$*. Suppose that there is a policy* π *achieves the optimal regret* $\widetilde{O}\left(T^{\frac{\alpha+1}{2\alpha+1}}\right)$ *over* $\mathbb{E}[d|p] = f(p) \in \mathcal{H}(\alpha, L(\alpha))$ *, then there exists a constant* $C > 0$ π *such that*

$$\sup_{f \in \mathcal{H}(\beta, L(\beta))} R^{\pi}(T) \geq \Omega\left(T^{\frac{\beta+1}{2\beta+1} + \frac{\beta(\alpha-\beta)}{2(2\beta+1)^2(2\alpha+1)}}\right).$$

Self-Similarity Condition

Definition

A function $g : [a, b] \to \mathbb{R}, [a, b] \subseteq [0, 1]$ is self-similar on $[a, b]$ with parameters $\beta, l \in \mathbb{Z}^+, M_1 \in \mathbb{R}_{\geq 0}, M_2 \in \mathbb{R}_+$ if for some positive integer $c > M_1$ it holds that

$$\max_{V \in \mathcal{V}_c} \sup_{p \in V} \left| \Gamma_l^V g(p) - g(p) \right| \geq M_2 \cdot 2^{-c\beta},$$

where we define

$$\mathcal{V}_c = \left\{ \left[a + \frac{i}{2^c}, a + \frac{i+1}{2^c} \right] \cap [0, 1], i = 0, 1, \ldots 2^c - 1 \right\}$$

for any positive integer c.

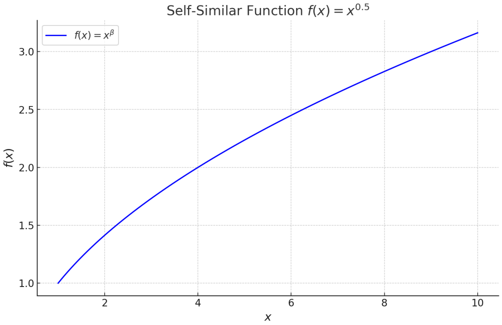
Examples of Self-Similar Functions

Example

Let f be any function with continuous first-order derivative uniformly bounded by C_1. We define the function class $\mathcal{F}$ as

$$\mathcal{F}(f) = \{f : x \mapsto c_0 \cdot x^{\beta} + f : c_0 \in \mathbb{R}, |c_0| \geq C_1\},$$

then all function in $\mathcal{F}(f)$ is self-similar with parameters $, l = 0$ for some constants M_1, M_2 depending on C_1 and C_2.



Self-Similarity Condition Doesn't Decrease Complexity

Theorem

For any positive parameters $\beta, M_1, L > 0$, *there exists a constant* $M_2 > 0$ *satisfying that*

$$\inf_{\pi} \sup_{f \in \mathcal{H}(\beta,L) \cap \mathcal{S}(\beta,w(\beta),M_1,M_2)} R^{\pi}(T) \geq \Omega(T^{\frac{\beta+1}{2\beta+1}}).$$

Self-Similarity Condition Doesn't Decrease Complexity

Theorem

For any positive parameters $\beta, M_1, L > 0$, there exists a constant $M_2 > 0$ satisfying that

$$\inf_{\pi} \sup_{f \in \mathcal{H}(\beta,L) \cap \mathcal{S}(\beta,w(\beta),M_1,M_2)} R^{\pi}(T) \geq \Omega(T^{\frac{\beta+1}{2\beta+1}}).$$

- Self-similarity does not change the minimax regret rate and therefore does not lower the problem difficulty for any admissible dynamic pricing policy π

Towards Smoothness-Adaptivity

Main idea behind learning smoothness parameter

- Employ two distinct levels of granularity to estimate the demand function, indexed by 1 and 2 respectively
- The difference of two estimates tells information on β

Lemma (Estimation Accuracy)

With $0 < \beta \leq \beta_{\max}$, we can obtain an estimator $\hat{\beta}$ in $2T^{\lfloor \frac{1}{2}+k_1 \rfloor}$ periods for $k_1 = \frac{1}{2\beta_{\max}+2}$, such that with probability at least $1 - O\left(e^{-C\ln^2(T)}\right)$,

$$\hat{\beta} \in \left[\beta - \frac{4\left(\beta_{\max}+1\right)\ln\left(\ln\left(T\right)\right)}{\ln\left(T\right)}, \beta\right].$$

Smoothness-Adaptive Dynamic Pricing Algorithm

Algorithm Smoothness-Adaptive Dynamic Pricing (`SADP`)

Input: Time horizon T, Hölder smoothness range $[\beta_{\min}, \beta_{\max}]$, minimum price $p_{\min}$, maximum demand $d_{\max}$;

Set $k_2 = \frac{1}{2\beta_{\max}+2}, k_2 = \frac{1}{4\beta_{\max}+2}$, $K_i = 2^{\lfloor k_i \log_2(T) \rfloor}$, for $i = 1, 2$;
for $i = 1, 2$ **do**
  Set trial time $T_i = T^{\lfloor \frac{1}{2} + k_i \rfloor}$;
  Fit local polynomial regression $\hat{f}_i$, respectively, on collected data from pulling T_i times uniformly selected prices;
end for
Let $\hat{\beta} = -\frac{\ln(\max \|\hat{f}_2 - \hat{f}_1\|_\infty)}{\ln(T)} - \frac{\ln(\ln(T))}{\ln(T)}$;
Call HSDP with $\hat{\beta}$;

Theorem

The cumulative regret of SADP is upper bounded by $\tilde{O}\left(T^{\frac{\beta+1}{2\beta+1}}\right)$.

Numerical Illustration

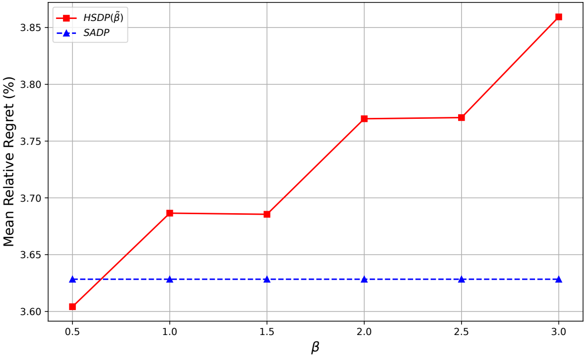


Figure: The comparison of mean relative regret over 30 experiments.

Summary

Takeaway

- To make the dynamic pricing algorithm practical, we need more adaptivity, but it is challenging in general

Our contributions

Summary

Takeaway

- To make the dynamic pricing algorithm practical, we need more adaptivity, but it is challenging in general

Our contributions

- We characterize adaptivity challenge

Summary

Takeaway

- To make the dynamic pricing algorithm practical, we need more adaptivity, but it is challenging in general

Our contributions

- We characterize adaptivity challenge
- We develop a smoothness-adaptive dynamic pricing algorithm under self-similarity

Summary

Takeaway

- To make the dynamic pricing algorithm practical, we need more adaptivity, but it is challenging in general

Our contributions

- We characterize adaptivity challenge
- We develop a smoothness-adaptive dynamic pricing algorithm under self-similarity
- We prove its optimal regret rate

Summary

Takeaway

- To make the dynamic pricing algorithm practical, we need more adaptivity, but it is challenging in general

Our contributions

- We characterize adaptivity challenge
- We develop a smoothness-adaptive dynamic pricing algorithm under self-similarity
- We prove its optimal regret rate

Paper available at arxiv.org/abs/2310.07558

Accepted to AISTATS 2024

Thank You!

Questions or comments? Contact: hansheng.jiang@rotman.utoronto.ca