Learning Algorithms for Dynamic Pricing: A Comparative Study

Chaitanya Amballa[1], **Narendhar Gugulothu**[1], **Manu K. Gupta**[2] and **Sanjay P. Bhat**[1]

[1]TCS Research and Innovation

[2]Indian Institute of Technology Roorkee

{chaitanya.amballa, narendhar.g, sanjay.bhat}@tcs.com
manu.gupta@ms.iitr.ac.in

Workshop on Real World Experiment Design and Active Learning, ICML 2020

July 18, 2020

Dynamic Pricing

- Maximize cumulative revenue over T periods by selecting a price p_t at each period t
- Revenue at period t is a noisy observation from a revenue function

$$r_t = g(p_t) + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \sigma^2) \tag{1}$$

- The revenue function is an unknown polynomial

$$g(p_t) = \tilde{\mu}_0 + \tilde{\mu}_1 p_t + \tilde{\mu}_2 p_t^2 + \cdots + \tilde{\mu}_n p_t^n$$

- Optimal revenue

$$r^* = \max_{p \in [p_{\min}, p_{\max}]} g(p)$$

- Cumulative regret

$$R(T) = \sum_{t=1}^{T} [r^* - \mathbb{E}(r_t)]$$

Goal

Objective

Learn the unknown parameters $\tilde{\mu}_0, \tilde{\mu}_1, \cdots, \tilde{\mu}_n$ from noisy observations of price and revenue pairs $\{(p_t, r_t)\}_{t=1}^{T}$ to suggest the optimal price while reducing the T-period expected cumulative regret, $R(T)$

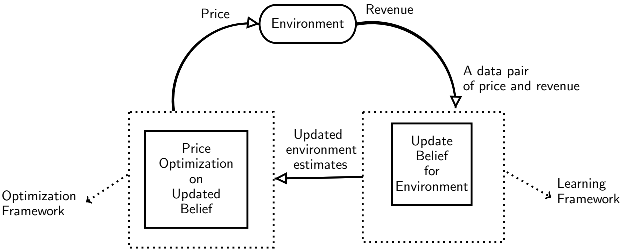

Figure: Dynamic pricing architecture

Algorithms Compared

- Standard Algorithms
  - Iterated leastsquare (ILS)
  - Constrained Iterated leastsquares (CILS)
  - Action Space Exploration (ASE)
  - Parameter Space Exploration (PSE)
  - Thompson Sampling (TS)
- Improved algorithms
  - Initial querying at Barycentric prices and doing a least squares fit
  - Controlled sampling by stopping criterion in TS
  - Controlled sampling by varying the exploration parameter σ in TS

Performance and Robustness Checks

- Regret performance for various degrees of the true revenue polynomial
- Robustness to mis-specification of the true degree
  - true polynomial degree $>$ assumed model degree
  - true polynomial degree $<$ assumed model degree
- Robustness to polynomial assumption