Data Science Lab
Workshop on Learning and Optimizing in Operations

Time: 12/10/2019, 9:00 am - 1:00 pm
Location: E25 - 401

Time	Speaker	Title
9:00 am - 9:15 am	David Simchi-Levi	Opening Remark
9:15 am - 10:00 am	Yining Wang	Multi-Modal Dynamic Pricing
10:00 am - 10:30 am	Yunzong Xu	Online Pricing with Offline Data: Phase Transition and Inverse Square Law
10:30 am - 10:45 am	Break	
10:45 am - 11:30am	Boxiao (Beryl) Chen	Tailored Base-Surge Policies in Dual-Sourcing Inventory Systems with Demand Learning
11:30 am - 12:15 pm	Zhenzhen Yan	A Representative Consumer Model in Data-Driven Pricing and Promotion Design Problem
12:15 pm - 12:45 pm	Ruihao Zhu	The Blessing of (Even More) Optimism: Reinforcement Learning under Non-Stationarity

Detailed Schedule

- 9:00 am - 9:15 am
  Speaker: Prof. David Simchi-Levi (MIT)
  Title: Opening remark

- 9:15 am- 10:00 am
  Speaker: Yining Wang (Univ. of Florida)
  Title: Multi-Modal Dynamic Pricing
  Abstract: We consider a stylistic question of dynamic pricing of a single product with demand learning. The candidate prices belong to a wide range of price interval, and the modeling of the demand functions is nonparametric in nature, imposing only smoothness regularity conditions. One important aspect of our modeling is the possibility of the expected reward function to be non-convex and indeed multi-modal, which leads to many conceptual and technical challenges. Our proposed algorithm is inspired by both the Upper-Confidence-Bound (UCB) algorithm for multi-armed bandit and the Optimism-in-Face-of-Uncertainty (OFU) principle arising from linear contextual bandits. Through rigorous regret analysis, we demonstrate that our proposed algorithm achieves optimal worst-case regret over a wide range of smooth function classes. More specifically, for k-times smooth functions and T selling periods, the regret of our propose algorithm is $\widetilde{O}(T^{(k+1)/(2k+1)})$, which is shown to be optimal via information theoretical lower bounds. We also show that in special cases such as strongly concave or infinitely smooth reward functions, our algorithm achieves an $O(\sqrt{T})$ regret matching optimal regret established in previous works. Finally, we present numerical results which verify the effectiveness of our method in numerical simulations.

- 10:00 am- 10:30 am
  Speaker: Yunzong Xu (MIT)
  Title: Online Pricing with Offline Data: Phase Transition and Inverse Square Law
  Abstract: This paper investigates the impact of pre-existing offline data on online learning, in the context of dynamic pricing. We study a single-product dynamic pricing problem over a selling horizon of T periods. The demand in each period is determined by the price of the product according to a linear demand model with unknown parameters. We assume that the seller already has some pre-existing offline data before the start of the selling horizon. The offline data set contains n samples, each of which is an input-output pair consists of a historical price and an associated demand observation. The seller wants to utilize both the pre-existing offline data and the sequential online data to minimize the regret of the online learning process.

  We characterize the joint effect of the size, location and dispersion of offline data on the optimal regret of the online learning process. Specifically, the size, location and dispersion of offline data are measured by the number of historical samples n, the absolute difference between the average historical price and the optimal price , and the standard deviation of the historical prices , respectively. We design a learning algorithm based on the "optimism in the face of uncertainty" principle, whose regret is optimal up to a logarithmic factor. Our results reveal surprising transformations of the optimal regret rate with respect to the size of offline data, which we refer to as phase transitions. In addition, our results demonstrate that the location and dispersion of the offline data set also have an intrinsic effect on the optimal regret, and we quantify this effect via the inverse-square law.

- 10:30 am- 10:45 am
  Break

- 10:45 am- 11:30 am
  Speaker: Boxiao (Beryl) Chen (UIC)
  Title: Tailored Base-Surge Policies in Dual-Sourcing Inventory Systems with Demand Learning
  Abstract: We consider a periodic-review dual-sourcing inventory system, in which the expedited supplier is faster and more costly, while the regular supplier is slower and cheaper. Under full demand distributional information, it is well-known that the optimal policy is extremely complex but the celebrated Tailored Base-Surge (TBS) policy performs near optimally. Under such a policy, a constant order is placed at the regular source in each period, while the order placed at the expedited source follows a simple order-up-to rule. In this paper, we assume that the firm does not know the demand distribution a priori, and makes adaptive inventory ordering decisions in each period based only on the past sales (a.k.a. censored demand) data. The standard performance measure is regret, which is the cost difference between a feasible learning algorithm and the clairvoyant (full-information) benchmark. When the benchmark is chosen to be the (full-information) optimal Tailored Base-Surge policy, we develop the first nonparametric learning algorithm that admits a regret bound of $O(T^{1/2}(logT)^3 \log\log T)$, which is provably tight up to a logarithmic factor. Leveraging the structure of this problem, our approach combines the power of bisection search and stochastic gradient descent and also involves a delicate high probability coupling argument between our and the clairvoyant optimal system dynamics. We also develop several technical results that are of independent interest.

- 11:30 am- 12:15 pm
  Speaker: Zhenzhen Yan (NTU)
  Title: A Representative Consumer Model in Data-Driven Pricing and Promotion Design Problem
  Abstract: We develop a data-driven approach to recover the "right" choice model for a

multi-product pricing problem, using the theory of a representative consumer in discrete choice. This approach uses a regularization function to capture diversification in choice behavior and establishes a set of closed-form relationships between the prices and choice probabilities with a separable function. By penalizing against deviation from these relationships in the data set, we propose a new loss function that is used to derive efficient algorithms for the inverse optimization problem, in both online and offline settings. This allows us to build tractable models for both estimation and price optimization problem. Extensive tests using both synthetic and industry data demonstrate the benefits of this approach in a multi-product pricing problem. By generating the representative consumer model to a sequential decision-making process, we further develop a data-driven framework to solve a bundle-pricing problem and a threshold-type of promotion optimization problem, which is widely used in e-commerce.

- 12:15 pm- 12:45 pm

  Speaker: Ruihao Zhu (MIT)

  Title: The Blessing of (Even More) Optimism: Reinforcement Learning under Non-Stationarity

  Abstract: We propose algorithms with state-of-the-art *dynamic regret* bounds for undiscounted reinforcement learning under drifting non-stationarity, where both the reward functions and state transition distributions are allowed to evolve over time as long as the total changes do not exceed certain *variation budgets*. We first develop a tuned Sliding Window Upper-Confidence bound for Reinforcement Learning with Confidence Widening (`SWUCRL2-CW`) algorithm, and show that it attains low dynamic regret bounds via a novel *budget-aware* analysis. Along the way, we identify a unique challenge associated with parameters estimation for reinforcement learning in gradually changing environments, and we demonstrate how a more optimistic design of learning algorithm can help to alleviate the issue. Finally, we present the Bandit-over-Reinforcement Learning (`BORL`) framework that further permits us to enjoy a parameter-free dynamic regret bound.