Almost Optimal Algorithms for Linear Stochastic Bandits with Heavy-Tailed Payoffs

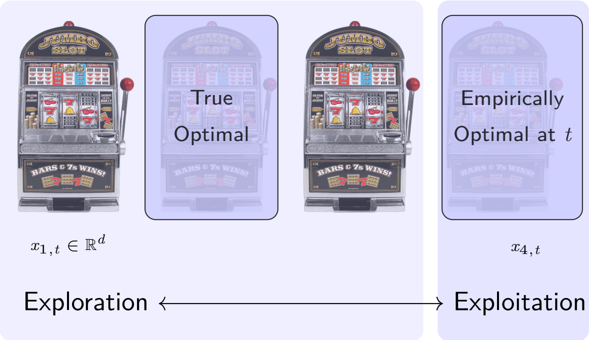
Han Shao*, Xiaotian Yu*, Irwin King and Michael R. Lyu

Department of Computer Science and Engineering
The Chinese University of Hong Kong

NeurIPS, Dec. 2018

Linear Stochastic Bandits (LSB)

Previous setting



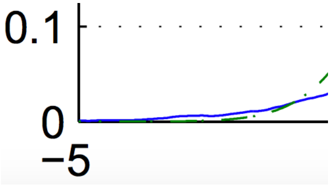
Learning setting

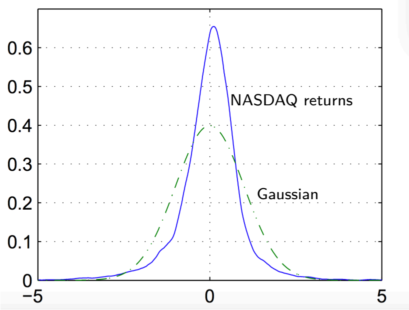
- 1. Given a set of arms represented by $D \subseteq \mathbb{R}^d$
- 2. At time t, select an arm $x_t \in D$, and observe $y_t(x_t) = \langle x_t, \theta_* \rangle + \eta_t$
- 3. The goal is to maximize $\sum_{t=1}^{T} \mathbb{E}[y_t(x_t)]$
- 4. η_t follows a sub-Gaussian distribution ($\mathbb{E}[\eta_t^2] < \infty$)

What Is A Heavy-Tailed Distribution?

Practical scenarios

- High-probability extreme returns in financial markets





- Many other real cases
  1. Delays in communication networks (Liebeherr et al., 2012)
  2. Analysis of biological data (Burnecki et al., 2015)
  3. ...

LSB with Heavy-Tailed Payoffs

Problem definition

- Multi-armed bandits (MAB) with heavy-tailed payoffs (Bubeck et al., 2013)

$$\mathbb{E}[\eta_t^{1+\epsilon}] < +\infty, \tag{1}$$

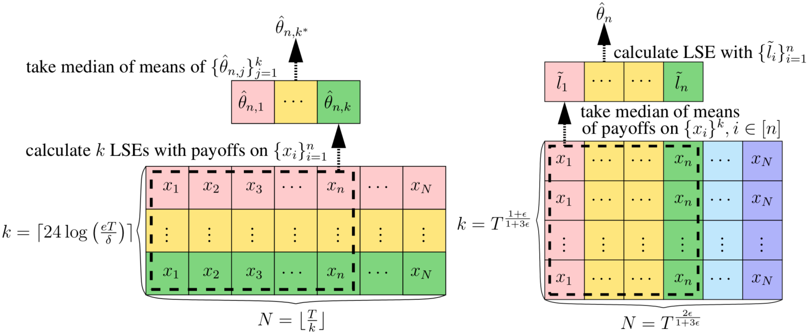
  where $\epsilon \in (0, 1]$
- Our setting: LSB with η_t satisfying Eq. (1)
  - Weaker assumption than sub-Gaussian
  - Medina and Yang (2016) studied LSB with heavy-tailed payoffs

	sub-Gaussian	heavy-tailed ($\epsilon = 1$)
MAB	$O(T^{\frac{1}{2}})$	$O(T^{\frac{1}{2}})$ by Bubeck et al. (2013)
LSB	$\widetilde{O}(T^{\frac{1}{2}})$	$\widetilde{O}(T^{\frac{3}{4}})$ by Medina and Yang (2016)

  - Can we achieve $\widetilde{O}(T^{\frac{1}{2}})$?

Algorithm: Median of means under OFU (MENU)

Framework comparison with MoM by Medina and Yang (2016)



(a) Framework of MENU (b) Framework of MoM

Regret Bounds

- Upper bounds

algorithm	MoM	MENU	CRT	TOFU
regret	$\widetilde{O}(T^{\frac{1+2\epsilon}{1+3\epsilon}})$	$\widetilde{O}(T^{\frac{1}{1+\epsilon}})$	$\widetilde{O}(T^{\frac{1}{2}+\frac{1}{2(1+\epsilon)}})$	$\widetilde{O}(T^{\frac{1}{1+\epsilon}})$

- Lower bound: $\Omega(T^{\frac{1}{1+\epsilon}})$

When $\epsilon = 1$, our algorithms achieve $\widetilde{O}(T^{\frac{1}{2}})$

See You at the Poster Session

Time: Dec. 5th, 10:45 AM – 12:45 PM
Location: Room 210 & 230 AB #158