PURDUE UNIVERSITY

Bootstrapping Upper Confidence Bound

Botao Hao, Yasin Abbasi-Yadkori, Zheng Wen, Guang Cheng

VinAI RESEARCH

DeepMind

Motivation

- Q: Can we have a **data-dependent**, **non-parametric**, **easy-to-implement** UCB algorithm?
  A: Yes, by multiplier bootstrap!
- Q: Can multiplier bootstrapped confidence bound ensure the **non-asymptotic validity**?
  A: Yes, by adding a **second-order correction!**
- **Warning!** Naive bootstrapped confidence bound $\to$ linear regret!

Bandits and Upper Confidence Bound

- **Multi-armed bandit as a showcase.**
  Pull an arm $I_t \in [K]$ and observes its reward y_{I_t} with an unknown mean μ_{I_t}.

$$\mathsf{REGRET}(T) = T\mu^* - \mathbb{E}\Big[\sum_{t=1}^T y_t\Big].$$

- **Upper confidence bound.**
  An upper confidence bound $\mathcal{G}(\boldsymbol{y}_n, 1-\alpha)$ for the true mean μ, of the form

$$\mathcal{G}(\boldsymbol{y}_n, 1-\alpha) = \big\{x \in \mathbb{R}, x - \bar{y}_n \le h_\alpha(\boldsymbol{y}_n)\big\},$$

where $\bar{y}_n$ is the **empirical mean**, $\alpha \in (0,1)$ is the confidence level, and $h_\alpha : \mathbb{R}^n \to \mathbb{R}^+$ is a threshold that could be either data-dependent (bootstrapped-based) or data-independent (concentration-based).

- **Non-asymptotics.**
  We define $\mathcal{G}(\boldsymbol{y}_n, 1-\alpha)$ as a non-asymptotic upper confidence bound if for **any sample size $n \ge 1$**, the following inequality holds

$$\mathbb{P}\Big(\mu \in \mathcal{G}(\boldsymbol{y}_n, 1-\alpha)\Big) \ge 1-\alpha.$$

Multiplier Bootstrap

- **Mean estimation.**
  Multiplier bootstrapped mean estimator:

$$\frac{1}{n}\sum_{i=1}^n w_i(y_i - \bar{y}_n) = \frac{1}{n}\sum_{i=1}^n (w_i - \bar{w}_n) y_i \overset{d}{\approx} \underbrace{\bar{y}_n - \mu}_{\text{target}}.$$

- **Bootstrap weights.**
  $\{w_i\}_{i=1}^n$ are some random variables independent of $\boldsymbol{y}_n$. Some classical weights are as follows:
  - Efron's bootstrap weights. $(w_1, \ldots, w_n)$ is a multinomial random vector with parameters $(n; n^{-1}, \ldots, n^{-1})$.
  - Gaussian weights. w_i's are i.i.d standard Gaussian random variables.
  - Rademacher weights. w_i's are i.i.d Rademacher variables.

Confidence Bound Based on Multiplier Bootstrap

- **Naive Bootstrap.**
  Approximate $(1-\alpha)$-quantile of $\bar{y}_n - \mu$ by $(1-\alpha)$-quantile of $n^{-1}\sum_{i=1}^n w_i(y_i - \bar{y}_n)$.
  The multiplier bootstrapped quantile is defined as,

$$q_\alpha(\boldsymbol{y}_n - \bar{y}_n) := \inf\Big\{x \in \mathbb{R} | \mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^n w_i(y_i - \bar{y}_n) > x\Big) \le \alpha\Big\}. \quad (1)$$

Question: if $q_\alpha(\boldsymbol{y}_n - \bar{y}_n)$ is a valid threshold for **any sample size** $n \ge 1$? **NO!**

- **Second-order Correction.**
  Classical statistical theories $\Rightarrow$ valid asymptotically ($n \to \infty$).
  Message: Valid **non-asymptotically** must pay the cost of a **second-order correction**.
- **Inform Theorem.**
  Require: symmetric random variables and Rademacher weights.
  For two arbitrary parameters $\alpha, \delta \in (0,1)$, the following inequality holds for any sample size $n \ge 1$,

$$\mathbb{P}_{\boldsymbol{y}}\Big(\bar{y}_n - \mu > \underbrace{q_{\alpha(1-\delta)}(\boldsymbol{y}_n - \bar{y}_n) + \sqrt{\frac{\log(2/\alpha\delta)}{n}}\varphi(\boldsymbol{y}_n)}_{\text{bootstrapped threshold}}\Big) \le 2\alpha, \quad (2)$$

where $\varphi(\boldsymbol{y}_n)$ is a non-negative function satisfying $\mathbb{P}_{\boldsymbol{y}}(|\bar{y}_n - \mu| \ge \varphi(\boldsymbol{y}_n)) \le \alpha$.
Special case for sub-Gaussian.

$$\text{bootstrapped threshold} = \underbrace{q_{\alpha/4}(\boldsymbol{y}_n - \bar{y}_n)}_{\text{main term}} + \underbrace{\frac{2\log(8/\alpha)}{n}}_{\text{second order correction}}.$$

Comparison between Confidence Bounds

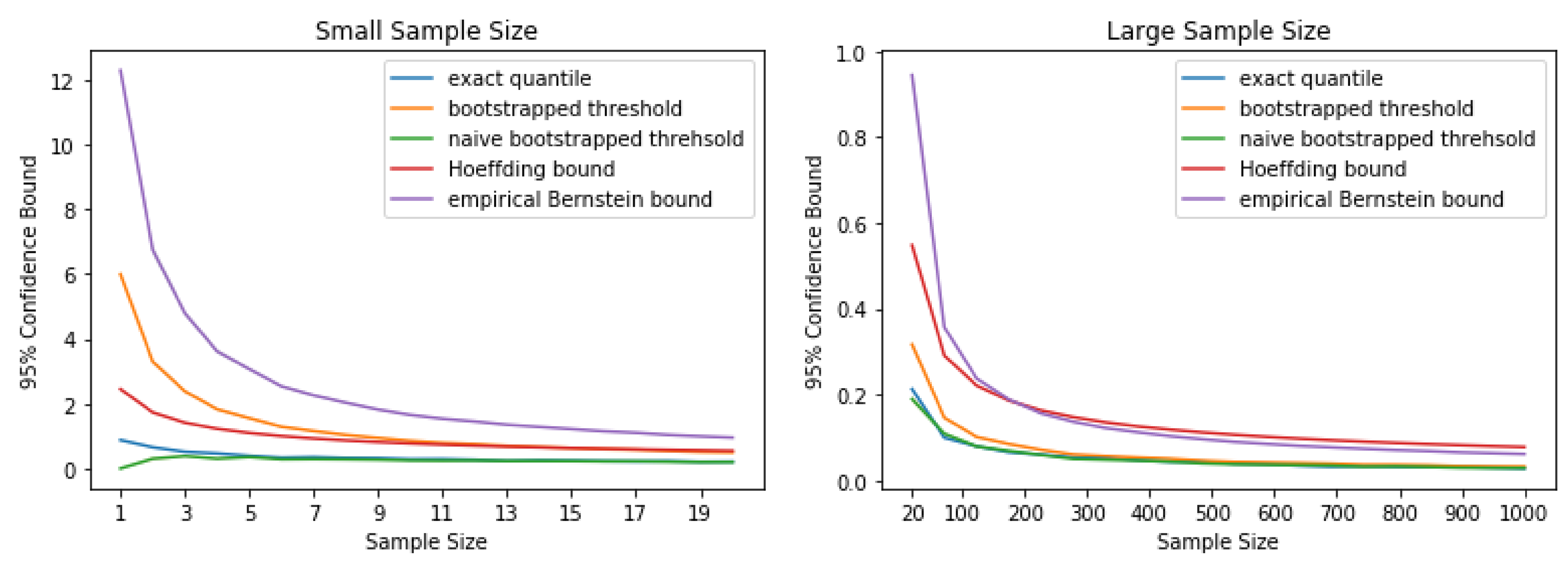

Figure 2: 95% confidence bound of the sample mean.

- Bootstrapped threshold **without correction is not valid** when sample size is small.
- Concentration-based threshold is **too loose**.

Regret Analysis

Definition 1 (Sub-Weibull Distribution)**.** We define y as a sub-Weibull random variable if it has a bounded ψ_β-norm. The ψ_β-norm of y for any $\beta > 0$ is defined as $\|y\|_{\psi_\beta} := \inf\{C \in (0,\infty) : \mathbb{E}[\exp(|y|^\beta / C^\beta)] \le 2\}$.

1. **Weaker assumption than sub-Gaussian or sub-exponential!**
2. $\beta = 2$**: sub-Gaussian;** $\beta = 1$**: sub-exponential**
3. **Novel concentration inequality derived.**

Theorem 0.1. *Consider a stochastic K-armed symmetric β-sub-Weibull bandit and let the confidence level $\alpha = 1/(t\log^2 t)$.*

- **Problem-dependent Regret**

$$R(T) \lesssim \sum_{k:\Delta_k > 0} \sigma^2 \frac{\log T}{\Delta_k} + \sigma K(\log T)^{1/\beta} + \sum_{k=2}^K \Delta_k.$$

- **Problem-independent Regret** *If the round $T \ge 2^{2/\beta - 3} K(\log T)^{2/\beta - 1}$,*

$$R(T) \lesssim \sigma\sqrt{TK\log T}. \quad (3)$$

Experiments

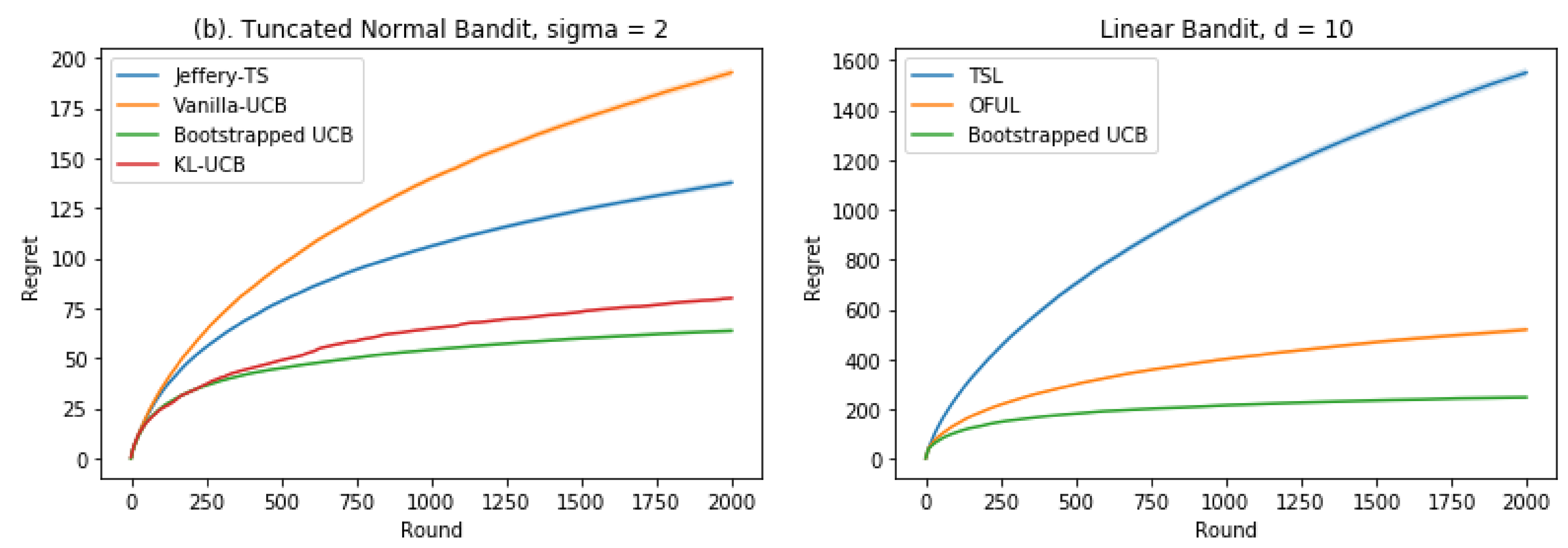

Figure 3: multi-armed bandit and linear bandit (without guarantee).

- TSL: Thompson sampling for linear bandit
- OFUL: optimism in the face of uncertainty linear bandit algorithm

Next?

- Relax symmetric assumption? Sharpen second-order correction term?
- Regret analysis for bootstrapped LinUCB? Bootstrap log-likelihood function...
- Extension to tabular MDP?