Note for Bandit Algorithms

Chenmi'en Tan
E-mail: chenmientan@outlook.com

Actively Updating (Last update: June 18, 2021)

Theorem 0.1

Suppose that $X - \mu$ obeys σ-subgaussian distribution, we have

$$\mathbb{P}(\hat{\mu} - \mu \geq \epsilon) \leq \exp(-\frac{n\epsilon^2}{2\sigma^2})$$

for any $\epsilon > 0$, where $\mu = \mathbb{E}[X]$ and $\hat{\mu} = \overline{X}$. Equivalently, for $\forall \delta \in (0, 1]$,

$$\mathbb{P}(\hat{\mu} - \mu \geq \sigma\sqrt{\frac{2}{n}\log\frac{1}{\delta}}) \leq \delta$$

1 Stochastic Bandits

Theorem 1.1: ETC

Suppose that $X_i - \mu_i$ obeys σ-subgaussian distribution, the regret for ETC is bounded by

$$m\sum_{i=1}^{k}\Delta_i + (n - mk)\sum_{i=1}^{k}\Delta_i \exp(-\frac{m\Delta_i^2}{4\sigma^2})$$

when $n \geq km$.

Theorem 1.2: UCB

Suppose that $X_i - \mu_i$ obeys σ-subgaussian distribution, the regret for UCB is bounded by

$$3\sum_{i=1}^{k}\Delta_i + \min\{\sum_{\Delta_i \neq 0}\frac{16\log n}{\Delta_i}, 8\sqrt{nk\log n}\}$$

when $\delta = 1/n^2$.

Theorem 1.3

Suppose that X obeys continuous distribution, then we have

$$\mathrm{CVaR}_{\alpha}(X) := \inf_{x\in\mathbb{R}}\{x + \frac{1}{\alpha}\mathbb{E}[(X-x)^{+}]\} = \mathbb{E}[X|X \geq \mathrm{VaR}_{\alpha}(X)]$$

Proof. Let $G(x) = x + \frac{1}{\alpha}\mathbb{E}[(X-x)^{+}]$. Notice that

$$\mathbb{E}[(X-x)^{+}] = \int_{F(x)}^{1}(t-x)\mathrm{d}F(t) = \int_{F(x)}^{1}(F^{-1}(u)-x)\mathrm{d}u$$

Then $\frac{\mathrm{d}G}{dx} = 1 - \frac{1}{\alpha}(1-F(x)) = 0 \Leftrightarrow x = F^{-1}(1-\alpha) = \mathrm{VaR}_{\alpha}(X)$ and $\frac{\mathrm{d}^2G}{\mathrm{d}x^2} = \frac{1}{\alpha}f(x) \geq 0$. Hence

$$\inf_{x\in\mathbb{R}} G(x) = G(\mathrm{VaR}_{\alpha}(X)) = \mathrm{VaR}_{\alpha}(X) + \underbrace{\frac{1}{\alpha}\mathbb{E}[(X-\mathrm{VaR}_{\alpha}(X))^{+}]}_{\mathbb{E}[X-\mathrm{VaR}_{\alpha}(X)|X\geq\mathrm{VaR}_{\alpha}(X)]} = \mathbb{E}[X|X \geq \mathrm{VaR}_{\alpha}(X)]$$

as we desire. $\square$

Theorem 1. *Suppose that* $\overline{R}_n = \sum_{t=1}^{n} \Delta_{A_t}$

Proof. For any $(u_1, \dots, u_k) \in \mathbb{N}_+^k$, define

$$G_i = \{\mu^* < \min_{t \in \{1,\dots,n\}} \mathrm{UCB}^*(t, \delta)\} \cap \{\hat{\mu}_{iu_i} + \sqrt{\frac{2}{u_i} \log \frac{1}{\delta}} < \mu^*\}$$

When G_i occurs, there is $T_i(n) \le u_i$. Hence $T_i(n) \ge u_i + 1$ implies G_i^c. Notice that

$$G_i^c \subset \{\mu^* \ge \min_{s \in \{1,\dots,n\}} \hat{\mu}^* + \sqrt{\frac{2}{s} \log \frac{1}{\delta}}\} \cup \{\hat{\mu}_{iu_i} - \mu_i \ge \Delta_i - \sqrt{\frac{2}{u_i} \log \frac{1}{\delta}}\}$$

Thus we have

$$\mathbb{P}(G_i^c) \le n\delta + \exp(-\frac{u_i(\Delta_i - \sqrt{\frac{2}{u_i} \log \frac{1}{\delta}})^2}{2})$$

There is

$$\mathbb{P}(\overline{R}_n \ge \sum_{\Delta_i \ne 0} (u_i + 1)\Delta_i) \le \mathbb{P}(\exists i \in \{1, \dots, k\}, \Delta_i \ne 0, T_i \ge u_i + 1) \le \sum_{\Delta_i \ne 0} \mathbb{P}(T_i \ge u_i + 1)$$

$$\le \sum_{\Delta_i \ne 0} \mathbb{P}(G_i^c) \le (k-1)n\delta + \sum_{\Delta_i \ne 0} \exp(-\frac{u_i(\Delta_i - \sqrt{\frac{2}{u_i} \log \frac{1}{\delta}})^2}{2})$$

$\square$

References