# Stochastic Linear Bandits with Finite Number of Arms.

**Source:** Chap 22, Bandit Algorithms, L & S. (text)

**Previous:** (Corollary 19.3 / Set 14): $R_n \leq C \cdot d \sqrt{n} \ln(n2)$

dimension

**Now:** ① Finite number of arms, $|\mathcal{A}| = k$, and arms are fixed over time

② Reward: $X_t = \langle A_t, \theta^* \rangle + \eta_t$, where

$$E\left[ e^{\phi \eta_t} \mid A_1, X_1, \dots, A_{t-1}, X_{t-1}, A_t \right] \leq e^{\phi^2/2} \quad a.s$$

(conditionally 1-subG)

③ $\Delta_a = \max_{b \in \mathcal{A}} \langle \theta^*, b - a \rangle \leq 1.$

i.e., all arms rewards are bounded wrt each other.

Algorithm: (Successive Elimination of Arms)

Initial: $\delta \in (0,1)$ $\mathcal{A} = \{A_1, \cdots, A_k\} \in \mathbb{R}^d$

$$\ell = 1, \ \mathcal{A}_1 = \mathcal{A}, \ t = 0, \ t_1 = 1$$

Iteration: ($\ell^{th}$ round)

ⓐ Choose $\pi_\ell^*$ a G-optimal design over $\mathcal{A}_\ell$.

ⓑ $$\varepsilon_\ell = \frac{1}{2^\ell} \quad \text{and} \quad N_\ell(a) = \left\lceil \frac{2d\, \pi_\ell(a)}{\varepsilon_\ell^2} \log\left(\frac{k\, \ell(\ell+1)}{2}\right) \right\rceil$$

$$\text{with } N_\ell = \sum_{a \in \mathcal{A}_\ell} N_\ell(a)$$

ⓒ Sample arms using (deterministic) $\{N_\ell(a)\}_{a \in \mathcal{A}_\ell}$.

ⓓ $$\hat{\theta}_\ell = V_\ell^{-1} \sum_{t=t_\ell}^{t_\ell + N_\ell} A_t X_t, \quad \text{and}$$

samples collected in round $\ell$.

$$V_\ell = \sum_{a \in \mathcal{A}_\ell} N_\ell(a) \cdot a a^T$$

(e) $\mathcal{A}_{\ell+1} = \left\{ a \in \mathcal{A}_\ell : \underbrace{\max_{b \in \mathcal{A}_\ell} \langle \hat{\theta}_\ell, b - a \rangle}_{} \leq 2\varepsilon_\ell \right\}$

keep only those arms which are scoring close to the best scoring current arm (recall elimination algorithm earlier).

$= \underbrace{\max_{b \in \mathcal{A}_\ell} \langle \hat{\theta}_\ell, b \rangle}_{\text{best scoring arm wrt current estimate}} - \langle \hat{\theta}_\ell, a \rangle$

(f) $t_{\ell+1} \leftarrow t_\ell + N_\ell$, $(\ell+1) \leftarrow \ell$

---

Theorem (22.1 in text): w.p. $\geq (1-\delta)$, for the algo. above,

$$R_n \leq C \sqrt{nd \ln\left(\frac{k \ln(n)}{\delta}\right)}$$

Further, with $\delta = 1/n$, $E[R_n] \leq C\sqrt{nd \ln(kn)}$

Pf: Follows from Set 16. $\boxed{}$