$(x_1, x_2)_2 = (-0.9748, -0.2424) = (n_1, n_2)_2$ $\qquad \hat{n}_2 \nparallel \hat{f}_2$

$(f_1, f_2)_2 = (0.49, 7.9583)$

Thus, it seems that this system can be simulated at the ellipse centred at the origin with equal axes.

EE6106 Lecture 9 (Date: 16 Feb 2024)

Recall: UCB was not an anytime algorithm, and required knowledge of the horizon. However, this is not a big flaw

Anytime UCB

$X_1 \dots X_n \sim$ iid 1-sub Gaussian with mean μ

$$\hat{\mu}_n = \frac{\sum_{1}^{n} x_i}{n}$$

$$\mathbb{P}\left(\mu \leq \hat{\mu}_n + \sqrt{\frac{2\log(1/\delta)}{n}}\right) \geq 1-\delta \quad \dashrightarrow \text{UCB}$$

$$\mathbb{P}\left(\mu \geq \hat{\mu}_n - \sqrt{\frac{2\log(1/\delta)}{n}}\right) \geq 1-\delta \quad \dashrightarrow \text{LCB}$$

Earlier, we used $\delta = \frac{1}{n^2}$ $\qquad$ (n was the horizon

Now, we use $\delta = t^{-4}$ $\qquad$ t is the current time instant)

Anytime UCB:

- Play each arm once
- At time $t > k$,

$$A_t = \arg\max_i \left[\hat{\mu}_i(t-1) + \sqrt{\frac{8\log t}{T_i(t-1)}}\right]$$

Thm Under this algo,

$$R_n \leq 5 \sum \Delta_i + 32 \log(n) \sum_{i:\Delta_i > 0} \frac{1}{\Delta_i}$$

Proof Fix suboptimal arm i. WLOG, arm 1 is optimal

Claim: At time $t > k$, fixed arm i is pulled only if one of the following holds –

① $\hat{\mu}_1(T_1(t-1)) \leq \mu_1 - \sqrt{\frac{8\log t}{T_1(t-1)}}$ ... (arm 1's UCB has been violated)

This will happen when arm 1 is underperforming

② $\hat{\mu}_i(T_i(t-1)) \geq \mu_i + \sqrt{\frac{8\log t}{T_i(t-1)}}$

This means that arm i is overperforming, i.e. the LCB of arm i is violated.

③ $T_i(t-1) \leq \frac{32\log t}{\Delta_i^2}$

This means that arm i has yet not been explored enough.

<u>Proof</u> :

Say ①, ② and ③ are violated

Then ①' $\Rightarrow$ UCB of arm 1 is valid $\Rightarrow \hat{\mu}_1 + \sqrt{\frac{8\log t}{T_i(t-1)}} > \mu_1$

$$\Rightarrow \hat{\mu}_1 + \sqrt{\frac{8\log t}{T_i(t-1)}} > \mu_i + \Delta_i > \Delta_i + \hat{\mu}_i - \sqrt{\frac{8\log t}{T_i(t-1)}}$$

The last inequality arises from the negation of ②, i.e. ②'

$$\Rightarrow \hat{\mu}_1 + \sqrt{\frac{8\log t}{T_i(t-1)}} > \underbrace{\left(\Delta_i - \sqrt{\frac{32\log t}{T_i(t-1)}}\right)}_{>0,\ \text{due to ③}'} + \hat{\mu}_i + \sqrt{\frac{8\log t}{T_i(t-1)}}$$

$\therefore$ UCB (arm 1) $>$ UCB (arm i)

$\square$

$\therefore$ If ①, ② and ③ are violated, then arm i cannot be pulled.

$$T_i(n) = \sum_1^n \mathbb{1}\{A_t = i\} \leq \sum_1^n \mathbb{1}\{①, \text{or } ②, \text{or } ③ \text{ holds at } t\}$$

$$\leq \sum_1^n \mathbb{1}\{① \text{ holds at } t\}$$

$$+ \sum_1^n \mathbb{1}\{② \text{ holds at } t\}$$

$$+ \sum_1^n \mathbb{1}\{③ \text{ holds at } t\}$$

$$\leq \sum_{1}^{n} \mathbb{1}_{\{① \text{ holds at } t\}} + \sum_{1}^{n} \mathbb{1}_{\{② \text{ holds at } t\}} + \left\lceil \frac{32 \log n}{\Delta_i^2} \right\rceil$$

$$\mathbb{E}[T_i(n)] \leq \sum_{1}^{n} \mathbb{P}_{\{① \text{ holds at } t\}} + \sum_{1}^{n} \mathbb{P}_{\{② \text{ holds at } t\}} + 1 + \frac{32 \log n}{\Delta_i^2}$$

$$\mathbb{P}(① \text{ holds at } t) \leq \mathbb{P}\left(\exists s \in [t] : \hat{\mu_1}(s) \leq \mu_1 - \sqrt{\frac{8 \log t}{s}} \right)$$

$$\leq \sum_{s=1}^{t} \mathbb{P}\left(\hat{\mu_1}(s) \leq \mu_1 - \sqrt{\frac{8 \log t}{s}}\right) \leq t\left(t^{-4}\right) \leq t^{-3}$$

$$\Rightarrow \sum_{1}^{n} \mathbb{P}(① \text{ holds at } t) \leq \sum_{1}^{n} t^{-3} \leq \sum_{1}^{\infty} t^{-3} \leq 2$$

What we did for arm 1 identically holds for arm i, so

$$\mathbb{P}(② \text{ holds at } t) \leq 2$$

$$\therefore \mathbb{E}[T_i(n)] \leq 5 + \frac{32 \log n}{\Delta_i^2}$$

$$\therefore \mathbb{E}[R_n] \leq 5 \sum \Delta_i + 32 \sum \frac{\log n}{\Delta_i}$$

□

Note the structure for each of these proofs. We try to establish some rare event and then apply bounds on the probabilities of these rare event.

Comments

① $\delta = t^{-4}$ is overkill.

② $\delta = t^{-\alpha}$ for $\alpha > 2$ works (look at the part where we compute $\mathbb{P}$(① holds at t))

③ It turns out that $\alpha > 1$ also works, in that, the algorithm will still give ~~algorithmic~~ logarithmic regrets, but the naïve approach used by us while defining the union bounds will break down.

④ The same approach will also work beyond sub-Gaussian distributions (such as Bernoulli). We shall need different type of concentration inequalities for heavy-tailed distributions (not Chernoff type)

Ref: to be informed

Lattimore covers various analyses for UCB variants

Consider two probability measures P and Q on some measurable space

$(\Omega, F) \longrightarrow$ Subsets of Ω to which we have decided to assign a probability

Ω $\rightarrow$ Sample Space

F $\rightarrow$ σ-algebra

Def^n:

$\mathbb{P}$ is absolutely continuous w.r.t $\mathbb{Q}$ if

$$\mathbb{Q}(A) = 0 \Rightarrow \mathbb{P}(A) = 0$$

$$\Updownarrow$$

$$\mathbb{P}(A) > 0 \Rightarrow \mathbb{Q}(A) > 0$$

Fact: If $\mathbb{P}$ is absolutely continuous wrt $\mathbb{Q}$, $\exists$ measurable f s.t. for any event A,

$$\mathbb{P}(A) = \int_A d\mathbb{P} = \int_A f \, d\mathbb{Q}$$

$f \rightarrow$ $\boxed{\text{RADON NIKODYM DERIVATIVE OF } \mathbb{P} \text{ wrt } \mathbb{Q}}$ $\rightarrow$ denoted $\frac{d\mathbb{P}}{d\mathbb{Q}}$

This means that we can integrate over $\mathbb{Q}$ instead of integrating over $\mathbb{P}$

implies

① $\mathbb{P}(A) = \mathbb{E}_{\mathbb{P}}[\mathbb{1}_A]$

$$= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_{(A)} \frac{d\mathbb{P}}{d\mathbb{Q}}\right]$$

For a random variable X,

$$\mathbb{E}_{\mathbb{P}}[X] = \mathbb{E}_{\mathbb{Q}}\left[X \frac{d\mathbb{P}}{d\mathbb{Q}}\right]$$

The Radon-Nikodym derivative is telling us how the relative likelihood under $\mathbb{P}$ is as compared to the relative likelihood under $\mathbb{Q}$

Eg: Suppose $\Omega = \{x_1, x_2, \dots, x_n\}$, where x_i are disjoint

$$\sigma(\Omega) = 2^{\Omega}$$

The full measure can be characterized by the probabilities of the singleton events.

Say $\mathbb{P}(\{x_i\}) = p_i \quad \left(\sum_1^n p_i = 1\right)$

(Outcome $\Rightarrow \{1, 2, 3, 4, 5, 6\}$ while Events $\Rightarrow$ Subsets)

Similarly define $\mathbb{Q}(\{x_i\}) = q_i$

$$\frac{d\mathbb{P}}{d\mathbb{Q}}(x_i) = \frac{p_i}{q_i} \quad \text{trivially}$$

$$\mathbb{P}(A) = \sum_{i : x_i \in A} p_i = \sum_{i : x_i \in A} q_i \frac{p_i}{q_i} = \sum_{i : x_i \in A} q_i \frac{d\mathbb{P}}{d\mathbb{Q}}(x_i)$$

$$= \mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_{(A)} \frac{d\mathbb{P}}{d\mathbb{Q}}\right]$$