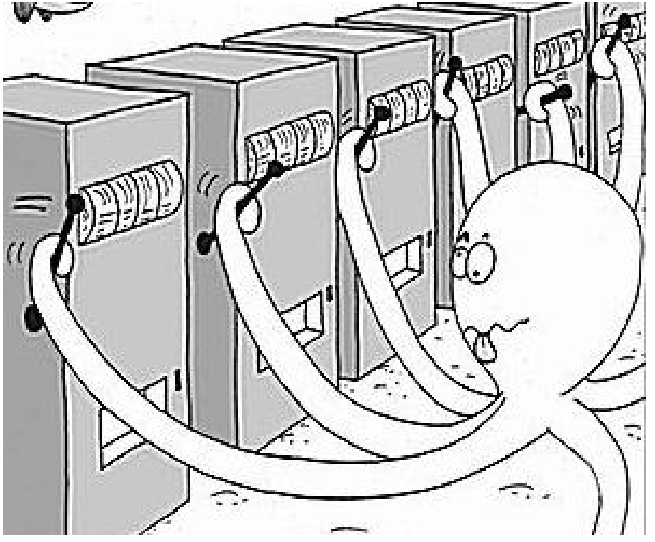
Stochastic Recursive Optimization: A Structured Multi-Armed Bandit Problem

Roland Szögi
Supervisor: Balázs Csanád Csáji

INTRODUCTION



Multi-Armed Bandit Model

- The multi-armed bandit model consists of a set of arms $\mathcal{A}$.
- To every arm $a \in \mathcal{A}$ belongs a distribution $\nu(a)$.
- In each round an arm $a \in \mathcal{A}$ is chosen and a reward $R(a)$ is sampled from distribution $\nu(a)$.
- An arm is called optimal, if it has the highest expected reward among all of the arms.

Definition 1.1

An arm $a \in \mathcal{A}$ is called ε-optimal if

$$\mathbb{E}[R(a)] \geq r^* - \varepsilon,$$

where r^* denotes the expectation of the optimal arm.

- The goal is to find an ε-optimal arm for a given ε.

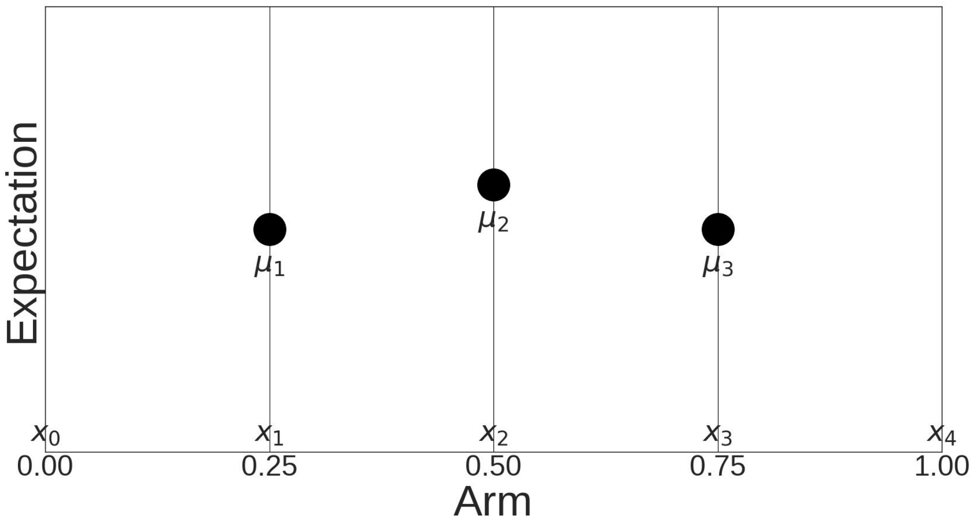
Definition 1.2

An algorithm is called (ε, δ)-PAC (probably approximately correct), if it returns an ε-optimal arm with probability at least $1 - \delta$.

Continuous Case

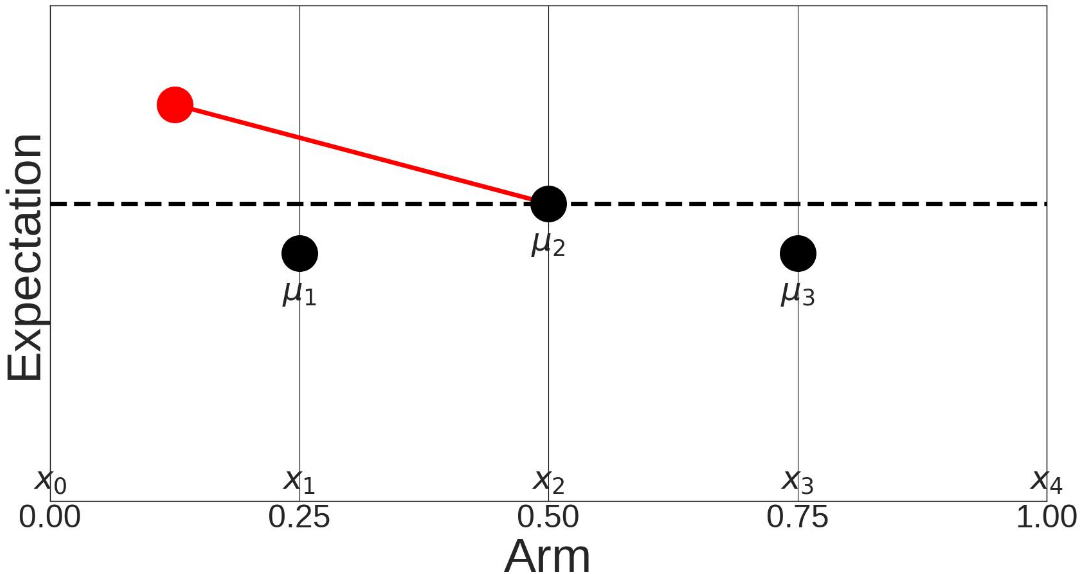
Assumption 1

The arms are the points of the $[0,1]$ *interval and the expectation of arm* $a \in [0,1]$ *is* $f(a)$*, where* $f : [0,1] \to \mathbb{R}$ *is an unknown concave function.*



OBSERVATIONS

$$\mu_1 \leq \mu_2 \implies f(x) \leq \mu_2 \quad \forall\, x \in [x_0, x_1].$$
$$\mu_3 \leq \mu_2 \implies f(x) \leq \mu_2 \quad \forall\, x \in [x_3, x_4].$$



Observations

$$\mu_1 \geq \mu_2 \implies f(x) \leq \mu_2 \quad \forall x \in [x_2, x_4].$$
$$\mu_3 \geq \mu_2 \implies f(x) \leq \mu_2 \quad \forall x \in [x_0, x_2].$$

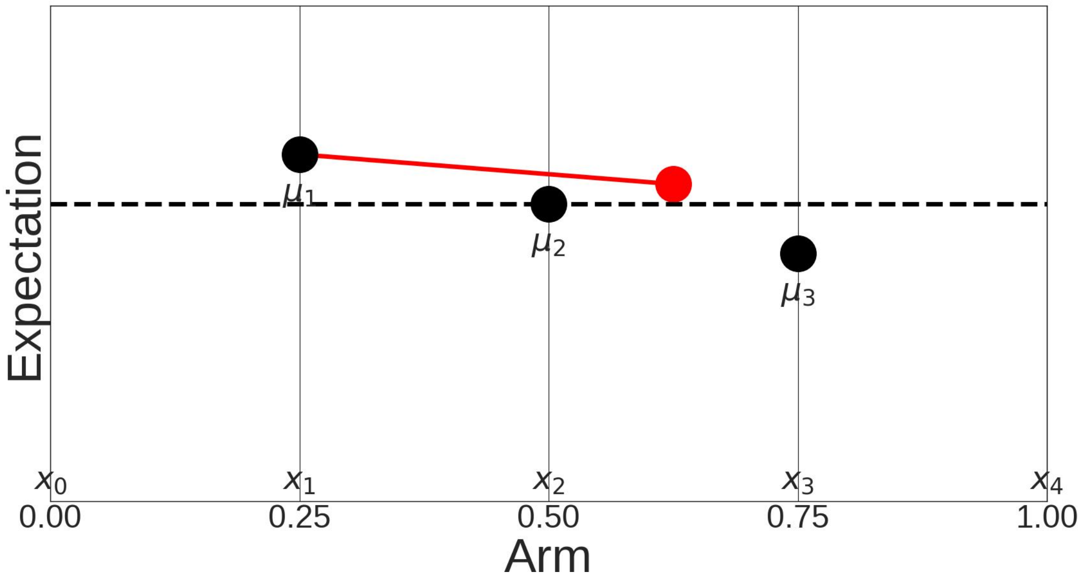
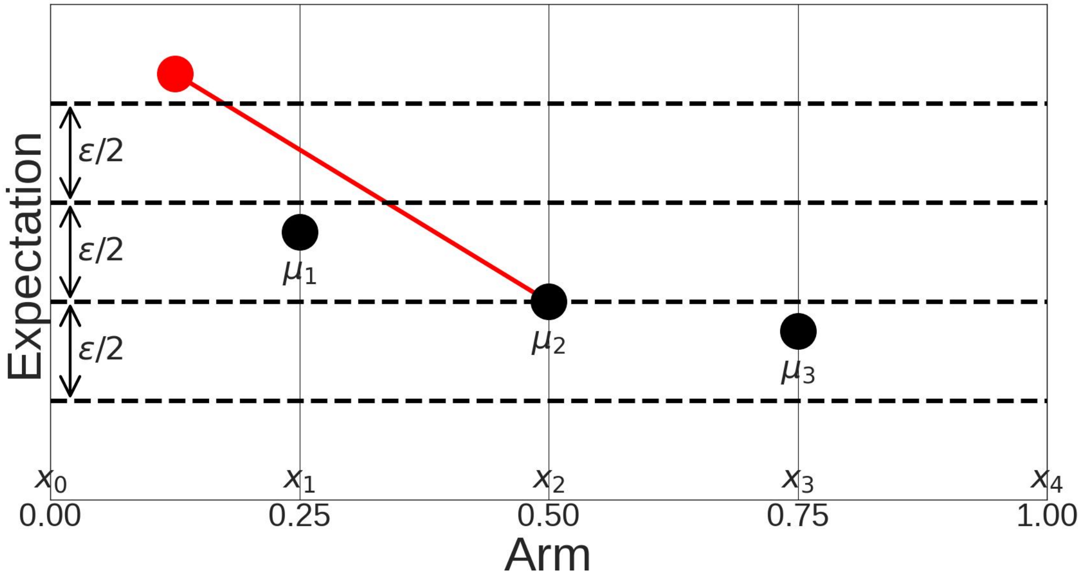


Figure. Enter Caption

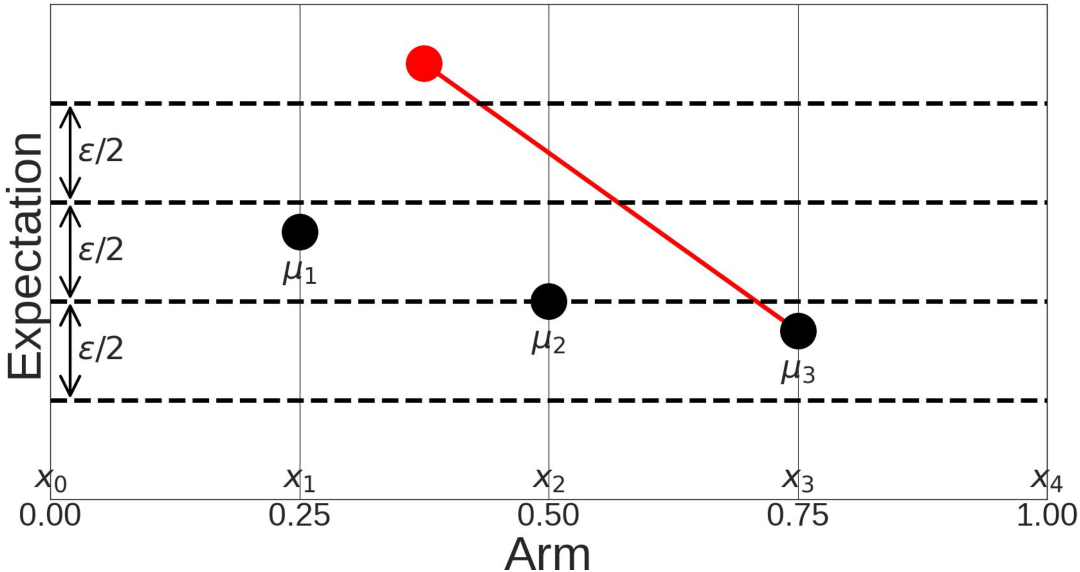
Observations

$$\mu_2 - \varepsilon/2 \le \mu_1, \mu_3 \le \mu_2 + \varepsilon/2 \implies f(x) \le \mu_2 + \varepsilon \quad \forall x \in [x_0, x_4].$$



Observations

$\mu_2 - \varepsilon/2 \leq \mu_1, \mu_3 \leq \mu_2 + \varepsilon/2 \implies f(x) \leq \mu_2 + \varepsilon \quad \forall x.$



Algorithm

If $|\mu_i - \hat{\mu}_i| \leq \varepsilon/8 \quad \forall i = 1,2,3$ then:

- $\hat{\mu}_1 \leq \hat{\mu}_2 - \varepsilon/4 \implies \mu_1 \leq \mu_2 \implies f(x) \leq \mu_2 \quad \forall x \in [x_0, x_1] \implies$ remove arms in $[x_0, x_1]$.
- $\hat{\mu}_1 \geq \hat{\mu}_2 + \varepsilon/4 \implies \mu_1 \geq \mu_2 \implies f(x) \leq \mu_2 \quad \forall x \in [x_2, x_4] \implies$ remove arms in $[x_2, x_4]$.
- $\hat{\mu}_3 \leq \hat{\mu}_2 - \varepsilon/4 \implies \mu_3 \leq \mu_2 \implies f(x) \leq \mu_2 \quad \forall x \in [x_3, x_4] \implies$ remove arms in $[x_3, x_4]$.
- $\hat{\mu}_3 \leq \hat{\mu}_2 + \varepsilon/4 \implies \mu_3 \geq \mu_2 \implies f(x) \leq \mu_2 \quad \forall x \in [x_0, x_2] \implies$ remove arms in $[x_0, x_2]$.
- $\hat{\mu}_2 - \varepsilon/4 \leq \hat{\mu}_1, \hat{\mu}_3 \leq \hat{\mu}_2 + \varepsilon/4 \implies \mu_2 - \varepsilon/2 \leq \mu_1, \mu_3 \leq \mu_2 + \varepsilon/2 \implies f(x) \leq \mu_2 + \varepsilon \quad \forall x$ $\implies$ return x_2.

We repeat this until the algorithm terminates.
We only make mistake when $\exists i \in \{1,2,3\} : |\hat{\mu} - \mu_i| \geq \varepsilon/8$. In round ℓ we sample the three arms so many times, that

$$\mathbb{P}(|\hat{\mu}_i - \mu_i| \geq \varepsilon/8) \leq \frac{\delta}{3 \cdot 2^\ell} \quad \forall i = 1,2,3.$$

In this case the probability of making a mistake in the ℓth round is

$$\mathbb{P}(\exists i \in \{1,2,3\} : |\hat{\mu} - \mu_i| \geq \varepsilon/8) \leq \frac{\delta}{2^\ell},$$

and the probability that the algorithm makes a mistake in any of the rounds is less than δ.

Subgaussian Random Variables

Definition 1.3

A random variable X is 1-subgaussian if for all $\lambda \in \mathbb{R}$:

$$\mathbb{E}[\exp(\lambda X)] \leq \exp\left(\frac{\lambda^2}{2}\right).$$

Statement 1

Assume that $X_i - \mu$ are independent, 1-subgaussian random variables. Then for any $\varepsilon > 0$,

$$\mathbb{P}(\hat{\mu} \geq \mu + \varepsilon) \leq \exp\left(-\frac{n\varepsilon^2}{2}\right),$$
$$\mathbb{P}(\hat{\mu} \leq \mu - \varepsilon) \leq \exp\left(-\frac{n\varepsilon^2}{2}\right)$$

where $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Remark 1

For random variables that are not centered ($\mathbb{E}[X] \neq 0$), the notation is abused by saying that X is 1-subgaussian if the noise $X - \mathbb{E}[X]$ is 1-subgaussian.

Algorithm

1. The arms are the points of the $[0,1]$ interval and the expectation of arm $a \in [0,1]$ is $f(a)$, where $f : [0,1] \to \mathbb{R}$ is an unknown concave function.
2. The arms are 1-subgaussian.

Input: $\delta > 0, \varepsilon > 0$

Output: an arm which is ε optimal with probability at least $1 - \delta$

Set $\ell = 1,\ \delta_1 = \delta/2,\ S_1 = [0,1],\ x_0^1 = 0,\ x_1^1 = 0.25,\ x_2^1 = 0.5,\ x_3^1 = 0.75,\ x_4^1 = 1.$

while TRUE **do**

$S_{\ell+1} = S_\ell.$

Sample $n_\ell = \lceil 128 \log(6/\delta_\ell)/\varepsilon^2 \rceil$ times the three arms: $x_1^\ell,\ x_2^\ell,\ x_3^\ell$.
Let $\hat{\mu}_1^\ell, \hat{\mu}_2^\ell, \hat{\mu}_3^\ell$ denote the sample means and $\mu_1^\ell, \mu_2^\ell, \mu_3^\ell$ denote the expectations.
if $\hat{\mu}_1^\ell, \hat{\mu}_3^\ell \in (\hat{\mu}_2^\ell - \varepsilon/4, \hat{\mu}_2^\ell + \varepsilon/4)$ **then return** x_2^ℓ.
if $\hat{\mu}_1^\ell \geq \hat{\mu}_2^\ell + \varepsilon/4$ **then** $S_{\ell+1} = S_{\ell+1} \setminus (x_2^\ell, x_4^\ell]$.
if $\hat{\mu}_1^\ell \leq \hat{\mu}_2^\ell - \varepsilon/4$ **then** $S_{\ell+1} = S_{\ell+1} \setminus [x_0^\ell, x_1^\ell)$.
if $\hat{\mu}_3^\ell \geq \hat{\mu}_2^\ell + \varepsilon/4$ **then** $S_{\ell+1} = S_{\ell+1} \setminus [x_0^\ell, x_2^\ell)$.
if $\hat{\mu}_3^\ell \leq \hat{\mu}_2^\ell - \varepsilon/4$ **then** $S_{\ell+1} = S_{\ell+1} \setminus (x_3^\ell, x_4^\ell]$.

$x_0^{\ell+1} = \min S_{\ell+1},\ x_4^{\ell+1} = \max S_{\ell+1},$
$x_1^{\ell+1} = \frac{3}{4} \cdot x_0^{\ell+1} + \frac{1}{4} \cdot x_4^{\ell+1},\ x_2^{\ell+1} = \frac{1}{2} \cdot x_0^{\ell+1} + \frac{1}{2} \cdot x_4^{\ell+1},\ x_3^{\ell+1} = \frac{1}{4} \cdot x_0^{\ell+1} + \frac{3}{4} \cdot x_4^{\ell+1}.$

$\delta_{\ell+1} = \delta_\ell/2,\ \ell = \ell + 1$

end while

Discrete Case

Assumption 2

There are n, arms numbered from 0 to $n-1$.

Assumption 3

The expectation of arm i is $f(i)$, where
$f : \mathbb{R} \to \mathbb{R}$ is an unknown concave function.

Assumption 4

The arms are 1-subgaussian.

Problem 1

We cannot use the previous algorithm, because the arms might not be divided into 4 equal-length subintervals with 3 arms.

Solution 1

Modify the algorithm so that the set of arms is halved in each round. With this algorithm we can solve the special case when there are $2^m + 1$ arms, and solve the general case with this special case.

Special Case

We want to remove half of the arms in each round.
The problematic cases, when the previous algorithm removes only one-fourth of the arms:

- $\hat{\mu}_1^\ell \in (\hat{\mu}_2^\ell - \varepsilon/4, \hat{\mu}_2^\ell + \varepsilon/4)$ and $\hat{\mu}_3^\ell \leq \hat{\mu}_2^\ell - \varepsilon/4$
- $\hat{\mu}_1^\ell \leq \hat{\mu}_2^\ell - \varepsilon/4$ and $\hat{\mu}_3^\ell \in (\hat{\mu}_2^\ell - \varepsilon/4, \hat{\mu}_2^\ell + \varepsilon/4)$

Solution 2

In the first case, sample the arm $x_{1.5}^\ell = (x_1^\ell + x_2^\ell)/2$ many times, estimate its expectation with the sample mean and remove arms based on this information. In the second case, sample $x_{2.5}^\ell = (x_2^\ell + x_3^\ell)/2$.

Special Case

If $\hat{\mu}_1^\ell \in (\hat{\mu}_2^\ell - \varepsilon/4, \hat{\mu}_2^\ell + \varepsilon/4)$ and $\hat{\mu}_3^\ell \leq \hat{\mu}_2^\ell - \varepsilon/4$, we sample $x_1^\ell, x_{1.5}^\ell, x_2^\ell$ so many times that with high probability $|\hat{\mu}_i^\ell - \mu_i^\ell| \leq \varepsilon/12 \quad \forall i = 1, 1.5, 2$.
If $|\hat{\mu}_i^\ell - \mu_i^\ell| \leq \varepsilon/12 \quad \forall i = 1, 1.5, 2$, then:

- $\hat{\mu}_1 \leq \hat{\mu}_{1.5} - \varepsilon/6 \implies \mu_1 \leq \mu_{1.5} \implies f(x) \leq \mu_{1.5} \quad \forall x \in [x_0, x_1] \implies$ remove arms in $[x_0, x_1]$.
- $\hat{\mu}_1 \geq \hat{\mu}_{1.5} + \varepsilon/6 \implies \mu_1 \geq \mu_{1.5} \implies f(x) \leq \mu_{1.5} \quad \forall x \in [x_{1.5}, x_4] \implies$ remove arms in $[x_{1.5}, x_4]$.
- $\hat{\mu}_2 \leq \hat{\mu}_{1.5} - \varepsilon/6 \implies \mu_2 \leq \mu_{1.5} \implies f(x) \leq \mu_{1.5} \quad \forall x \in [x_2, x_4] \implies$ remove arms in $[x_2, x_4]$.
- $\hat{\mu}_2 \leq \hat{\mu}_{1.5} + \varepsilon/6 \implies \mu_2 \geq \mu_{1.5} \implies f(x) \leq \mu_{1.5} \quad \forall x \in [x_0, x_{1.5}] \implies$ remove arms in $[x_0, x_{1.5}]$.
- $\hat{\mu}_{1.5} - \varepsilon/6 \leq \hat{\mu}_1, \hat{\mu}_2 \leq \hat{\mu}_{1.5} + \varepsilon/6 \implies \mu_{1.5} - \varepsilon/3 \leq \mu_1, \mu_2 \leq \mu_{1.5} + \varepsilon/3 \implies f(x) < \mu_{1.5} + \varepsilon \quad \forall x$ $\implies$ return $x_{1.5}$.

General Case

If there are $2^m + 1 < n < 2^{m+1} + 1$ arms, we can do the following:
Run the first round of the new algorithm with arms

$$\left\{i : \left\lfloor \frac{n}{2} \right\rfloor - 2^{m-1} \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor + 2^{m-1}\right\},$$
$$x_0^1 = \left\lfloor \frac{n}{2} \right\rfloor - 2^{m-1},\ x_1^1 = \left\lfloor \frac{n}{2} \right\rfloor - 2^{m-2},\ x_2^1 = \left\lfloor \frac{n}{2} \right\rfloor,$$
$$x_3^1 = \left\lfloor \frac{n}{2} \right\rfloor + 2^{m-2},\ x_4^1 = \left\lfloor \frac{n}{2} \right\rfloor + 2^{m-1}$$

and $\delta_0 = \delta/2$. If after the first round:

- $S_2 = \{i : x_0^1 \leq i \leq x_2^1\}$, then set
$$S = \{i : 0 \leq i \leq 2^m\},$$
- $S_2 = \{i : x_1^1 \leq i \leq x_3^1\}$, then set
$$S = \left\{i : \left\lfloor \frac{n}{2} \right\rfloor - 2^{m-1} \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor + 2^{m-1}\right\},$$
- $S_2 = \{i : x_2^1 \leq i \leq x_4^1\}$, then set
$$S = \{i : n - 2^m \leq i \leq n\}.$$

The length of S is $2^m + 1$ in all cases. Run the new algorithm with S and $\delta_0 = \delta/8$.

Sample Complexity

Theorem 1

The sample complexity of this algorithm when there are n arms:

$$\mathcal{O}\left(\frac{1}{\varepsilon^2}\left((\log n)^2 + \log n \cdot \log \frac{1}{\delta}\right)\right).$$

In this case the Median Elimination algorithm could be used to find an ε-optimal arm with probability at least $1 - \delta$ with a sample complexity of

$$\mathcal{O}\left(\frac{n}{\varepsilon^2} \log \frac{1}{\delta}\right).$$

Sample Complexity

If there is a known Δ such that $|\mu_i - \mu_{i-1}| \le \Delta$, $i = 1, 2, ..., n$, then the algorithm can terminate when the number of arms is $2 \cdot \lfloor \frac{\varepsilon}{\Delta} \rfloor + 1$ or less, by returning the arm in the middle. In this case the algorithm can terminate after $\left\lfloor \log_2 \frac{n}{2\lfloor \varepsilon/\Delta \rfloor + 1} \right\rfloor$ rounds so the sample complexity in this case is:

$$\mathcal{O}\left(\frac{\ell^2}{\varepsilon^2} + \frac{\ell}{\varepsilon^2} \log \frac{1}{\delta} \right).$$

where $\ell = \left\lfloor \log_2 \frac{n}{2\lfloor \varepsilon/\Delta \rfloor + 1} \right\rfloor$.
Similarly, if f is Lipschitz continuous with Lipschitz constant L, then the algorithm can terminate, when the length of the interval is less than or equal to $2 \cdot \varepsilon/L$ by returning the arm in the middle. In this case the algorithm terminates in

$$\ell = \left\lfloor \log_2 \frac{L}{\varepsilon} \right\rfloor$$

rounds so the sample complexity is

$$\mathcal{O}\left(\frac{1}{\varepsilon^2} \left(\left(\log \frac{L}{\varepsilon} \right)^2 + \left(\log \frac{L}{\varepsilon} \right) \cdot \log \frac{1}{\delta} \right) \right).$$

.

Thank You For Your Attention!