Stochastic Recursive Optimization: A Structured Multi-Armed Bandit Problem

Roland Szögi
Supervisor: Balázs Csanád Csáji

Multi-Armed Bandits

- 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.
- The goal is to find an ε-optimal arm.

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.

FURTHER ASSUMPTIONS

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.

Definition 2.1

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

$$\mathbb{E}[\exp(\lambda X)] \le \exp(\lambda^2\sigma^2/2).$$

Statement 1

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

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

$$\mathbb{P}(\hat{\mu} \le \mu - \varepsilon) \le \exp\left(-\frac{n\varepsilon^2}{2\sigma^2}\right)$$

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

Algorithm

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

Thank You For Your Attention!