Advanced Algorithms (IV)

Shanghai Jiao Tong University

Chihao Zhang

March 23rd, 2020

Review

Review

We learnt the Markov inequality

$$\Pr[X \geq a] \leq \frac{\mathbf{E}[X]}{a}$$

Review

We learnt the Markov inequality

$$\Pr[X \geq a] \leq \frac{\mathbf{E}[X]}{a}$$

We can choose an increasing function f so that

$$\Pr[X \geq a] = \Pr[f(X) \geq f(a)] \leq \frac{\mathbf{E}[f(X)]}{f(a)}$$

$f(x) = x^2$ yields the Chebyshev's inequality

$f(x) = x^2$ yields the Chebyshev's inequality

$$\Pr[\,|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathbf{Var}[X]}{a^2} = \frac{\mathbf{E}[X^2] - \mathbf{E}[X]^2}{a^2}$$

$f(x) = x^2$ yields the Chebyshev's inequality

$$\Pr[\,|X - \mathbf{E}[X]|\, \geq a] \leq \frac{\mathbf{Var}[X]}{a^2} = \frac{\mathbf{E}[X^2] - \mathbf{E}[X]^2}{a^2}$$

What is a good choice of f ?

$f(x) = x^2$ yields the Chebyshev's inequality

$$\Pr[\,|X - \mathbf{E}[X]|\, \geq a] \leq \frac{\mathbf{Var}[X]}{a^2} = \frac{\mathbf{E}[X^2] - \mathbf{E}[X]^2}{a^2}$$

What is a good choice of f ?

- f grows fast
- $\mathbf{E}[f(X)]$ is bounded and easy to calculate

Moment Generating Function

Moment Generating Function

The function $f(x) = e^{tx}$ is a natural choice

Moment Generating Function

The function $f(x) = e^{tx}$ is a natural choice

The function $\mathbf{E}[f(X)] = \mathbf{E}[e^{tX}]$ is called the *moment generating function*

Moment Generating Function

The function $f(x) = e^{tx}$ is a natural choice

The function $\mathbf{E}[f(X)] = \mathbf{E}[e^{tX}]$ is called the *moment generating function*

In some cases, $\mathbf{E}[e^{tX}]$ is easy to calculate…

Chernoff Bound

Chernoff Bound

Assume $X = \sum_{i=1}^{n} X_i$, where each $X_i \sim \text{Ber}(p_i)$ is an independent Bernoulli variable with mean p_i

Chernoff Bound

Assume $X = \sum_{i=1}^{n} X_i$, where each $X_i \sim \text{Ber}(p_i)$ is an independent Bernoulli variable with mean p_i

Chernoff Bound

Assume $X = \sum_{i=1}^{n} X_i$, where each $X_i \sim \text{Ber}(p_i)$ is an independent Bernoulli variable with mean p_i

---

$$
\begin{aligned}
\mathbf{E}[e^{tX}] &= \mathbf{E}[e^{t\sum_{i=1}^{n} X_i}] \\
&= \prod_{i=1}^{n} \mathbf{E}[e^{X_i}] = \prod_{i=1}^{n} \left(p_i \cdot e^t + 1 - p_i\right) \\
&= \prod_{i=1}^{n} e^{p_i(e^t-1)} = e^{\mathbf{E}[X](e^t-1)}
\end{aligned}
$$

Let $\mu = \mathbf{E}[X] = \sum_{i=1}^{n} p_i$

Let $\mu = \mathbf{E}[X] = \sum_{i=1}^{n} p_i$

For $t > 0$, we can deduce

Let $\mu = \mathbf{E}[X] = \sum_{i=1}^{n} p_i$

For $t > 0$, we can deduce

$$\Pr[X > (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}]$$

$$\leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1+\delta)\mu}} = \boxed{\frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} \quad \star}$$

Let $\mu = \mathbf{E}[X] = \sum_{i=1}^{n} p_i$

For $t > 0$, we can deduce

$$\Pr[X > (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}]$$
$$\leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1+\delta)\mu}} = \boxed{\frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} \star}$$

In order to obtain a tight bound, we optimize t to minimize $\star$

Since $\dfrac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = e^{\mu(e^t-1-t(1+\delta))}$, we can choose $t = \log(1+\delta) > 0$.

Since $\dfrac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = e^{\mu(e^t-1-t(1+\delta))}$, we can choose $t = \log(1+\delta) > 0$.

So $\Pr[X > (1+\delta)\mu] \leq \left(\dfrac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$

Since $\dfrac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = e^{\mu(e^t-1-t(1+\delta))}$, we can choose $t = \log(1+\delta) > 0$.

So $\Pr[X > (1+\delta)\mu] \le \left(\dfrac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$

We can similarly obtain (using $t < 0$)

Since $\dfrac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = e^{\mu(e^t-1-t(1+\delta))}$, we can choose $t = \log(1+\delta) > 0$.

So $\Pr[X > (1+\delta)\mu] \leq \left(\dfrac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$

We can similarly obtain (using $t < 0$)

$$\Pr[X < (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$$

To summarize, for $X = \sum_{i=1}^{n} X_i$, we have

- $\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$

- $\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$

To summarize, for $X = \sum_{i=1}^{n} X_i$, we have

- $\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$
- $\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$

A more useful expression is that for $0 < \delta \leq 1$

- $\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/3}$
- $\Pr[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}$

Max Load

Max Load

Recall in the max load problem, we throw n balls into n bins

Max Load

Recall in the max load problem, we throw n balls into n bins

The number of balls in i-th bin, $X_i \sim \text{Bin}\left(n, \frac{1}{n}\right)$

Max Load

Recall in the max load problem, we throw n balls into n bins

The number of balls in i-th bin, $X_i \sim \text{Bin}\left(n, \frac{1}{n}\right)$

Note that $\mathbf{E}[X_i] = 1$, what is the probability that $X_i > \dfrac{c \log n}{\log \log n}$?

In this case, $1 + \delta = \dfrac{c \log n}{\log \log n}$.

In this case, $1 + \delta = \dfrac{c \log n}{\log \log n}$.

Applying Chernoff bound, we obtain

In this case, $1 + \delta = \dfrac{c \log n}{\log \log n}$.

Applying Chernoff bound, we obtain

$$\Pr[X_i \geq \frac{c \log n}{\log \log n}] \leq \frac{e^{\delta}}{(1+\delta)^{1+\delta}} \leq n^{-c+o(1)},$$

In this case, $1 + \delta = \dfrac{c \log n}{\log \log n}$.

Applying Chernoff bound, we obtain

$$\Pr[X_i \geq \frac{c \log n}{\log \log n}] \leq \frac{e^{\delta}}{(1+\delta)^{1+\delta}} \leq n^{-c+o(1)},$$

which is tight in order comparing to our analytic result.

The Chernoff bound has a few drawbacks:

The Chernoff bound has a few drawbacks:

- each X_i needs to be *independent*.
- X_i is required to follow the $\mathrm{Ber}(p_i)$

The Chernoff bound has a few drawbacks:

- each X_i needs to be *independent*.
- X_i is required to follow the $\mathrm{Ber}(p_i)$

We will try to generalize the Chernoff bound to overcome these issues

Hoeffding Inequality

Hoeffding Inequality

The Hoeffding Inequality generalizes to those X_i with $\mathbf{E}[X_i] = 0$ and $a_i \leq X_i \leq b_i$.

Hoeffding Inequality

The Hoeffding Inequality generalizes to those X_i with $\mathbf{E}[X_i] = 0$ and $a_i \leq X_i \leq b_i$.

$$\Pr\left[\sum_{i=1}^{n} X_i \geq t\right] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

The key property to establish Hoeffding inequality is an upper bound on the moment generating function

The key property to establish Hoeffding inequality is an upper bound on the moment generating function

Lemma

Assume X satisfies $X \in [a, b]$ and $\mathbf{E}[X] = 0$, then

$$\mathbf{E}[e^{tX}] \leq \exp\left(\frac{t^2}{8}(b-a)^2\right)$$

The key property to establish Hoeffding inequality is an upper bound on the moment generating function

Lemma

Assume X satisfies $X \in [a, b]$ and $\mathbf{E}[X] = 0$, then

$$\mathbf{E}[e^{tX}] \leq \exp\left(\frac{t^2}{8}(b-a)^2\right)$$

You can find the proof of the lemma and Hoeffding inequality in the book *Probability and Computing*

Multi-Armed Bandit

Multi-Armed Bandit

In the problem of MAB, there are k bandits

Multi-Armed Bandit

In the problem of MAB, there are k bandits

- each bandit has a *unknown* *random* reward distribution f_i on $[0,1]$ with $\mu_i = \mathbf{E}[f_i]$
- each round one can pull an arm i and obtain a reward $r \sim f_i$

Multi-Armed Bandit

In the problem of MAB, there are k bandits

- each bandit has a *unknown* *random* reward distribution f_i on $[0,1]$ with $\mu_i = \mathbf{E}[f_i]$
- each round one can pull an arm i and obtain a reward $r \sim f_i$

The goal is to identify the best arm via trials

Regret of MAB

Regret of MAB

We assume $\mu_1 = \max_{1 \leq i \leq k} \mu_i$

Regret of MAB

We assume $\mu_1 = \max_{1 \leq i \leq k} \mu_i$

If the game is played for T rounds, the best reward on can obtain is $T\mu_1$ *in expectation*

Regret of MAB

We assume $\mu_1 = \max_{1 \le i \le k} \mu_i$

If the game is played for T rounds, the best reward on can obtain is $T\mu_1$ *in expectation*

We are often not so lucky to achieve this, so the goal is to find a strategy to minimize

Regret of MAB

We assume $\mu_1 = \max_{1 \le i \le k} \mu_i$

If the game is played for T rounds, the best reward on can obtain is $T\mu_1$ *in expectation*

We are often not so lucky to achieve this, so the goal is to find a strategy to minimize

$$R(T) = T\mu_1 - \sum_{t=1}^{T} \mu_{a_t}$$

Regret

Best Reward

a_t - the arm actually pulled at round t

What is a good strategy?

What is a good strategy?

We view $R(T)$ as a function of T and consider $T \to \infty$

What is a good strategy?

We view $R(T)$ as a function of T and consider $T \to \infty$

If we eventually find the best arm, then $R(T) = o(T)$

What is a good strategy?

We view $R(T)$ as a function of T and consider $T \to \infty$

If we eventually find the best arm, then $R(T) = o(T)$

If we fail to find the best arm, we will suffer a regret $\Omega(\Delta T)$, where Δ the gap between the optimal and suboptimal rewards

What is a good strategy?

We view $R(T)$ as a function of T and consider $T \to \infty$

If we eventually find the best arm, then $R(T) = o(T)$

If we fail to find the best arm, we will suffer a regret $\Omega(\Delta T)$, where Δ the gap between the optimal and suboptimal rewards

So we need the failure probability is $O(1/T)$

The Upper Confidence Bound Algorithm

The Upper Confidence Bound Algorithm

We collect information up to round T

The Upper Confidence Bound Algorithm

We collect information up to round T

- $n_i(T)$ - number of times that i-th arm has been pulled

- $\hat{\mu}_i(T)$ - estimate of the mean μ_i, which is equal to $\dfrac{\sum_{t=1}^{T} \mathbf{1}[a_t = i] \cdot r(t)}{n_i(T)}$ if $n_i(T) \neq 0$ and $r(t)$ is the reward at t-th round

Choose the Best Arm So Far?

Choose the Best Arm So Far?

The most straightforward idea is to choose the arm with best $\hat{\mu}_i(T)$

Choose the Best Arm So Far?

The most straightforward idea is to choose the arm with best $\hat{\mu}_i(T)$

The strategy might be inferior in case that we are unlucky so that the best arm performs bad at the first few trials.

Choose the Best Arm So Far?

The most straightforward idea is to choose the arm with best $\hat{\mu}_i(T)$

The strategy might be inferior in case that we are unlucky so that the best arm performs bad at the first few trials.

So we have to add some offset term for those arms that are not “well-explored”

The UCB algorithm chooses the arm with largest

$$\hat{\mu}_i(T) + c_i(T)$$

$c_i(T)$ - the confidence term of arm i at round T

The UCB algorithm chooses the arm with largest

$$\hat{\mu}_i(T) + c_i(T)$$

$c_i(T)$ - the confidence term of arm i at round T

The UCB algorithm chooses the arm with largest

$$\hat{\mu}_i(T) + c_i(T)$$

$c_i(T)$ - the confidence term of arm i at round T

Intuitively, $c_i(T)$ should be decreasing in n_i, so we give more chances to arms that have not been well-tested

The UCB algorithm chooses the arm with largest

$$\hat{\mu}_i(T) + c_i(T)$$

$c_i(T)$ - the confidence term of arm i at round T

Intuitively, $c_i(T)$ should be decreasing in n_i, so we give more chances to arms that have not been well-tested

Let's find out how to set $c_i(T)$

We need the following event to happen whp

$$\forall 2 \le i \le k, \quad \hat{\mu}_1(T) + c_1(T) > \hat{\mu}_i(T) + c_i(T)$$

We need the following event to happen whp

$$\forall 2 \leq i \leq k, \quad \hat{\mu}_1(T) + c_1(T) > \hat{\mu}_i(T) + c_i(T)$$

A sufficient condition for this is

$$\hat{\mu}_1(T) + c_1(T) > \mu_1 > \mu_i + 2c_i(T) > \hat{\mu}_i(T) + c_i(T)$$

We need the following event to happen whp

$$\forall 2 \le i \le k, \quad \hat{\mu}_1(T) + c_1(T) > \hat{\mu}_i(T) + c_i(T)$$

A sufficient condition for this is

$$\hat{\mu}_1(T) + c_1(T) \boxed{>} \mu_1 \boxed{>} \mu_i + 2c_i(T) \boxed{>} \hat{\mu}_i(T) + c_i(T)$$

$$\boxed{>} + \boxed{>} : \forall j, \quad \hat{\mu}_j(T) - c_j(T) < \mu_j < \hat{\mu}_j(T) + c_j(T)$$

We need the following event to happen whp

$$\forall 2 \le i \le k, \quad \hat{\mu}_1(T) + c_1(T) > \hat{\mu}_i(T) + c_i(T)$$

A sufficient condition for this is

$$\hat{\mu}_1(T) + c_1(T) \boxed{>} \mu_1 \boxed{>} \mu_i + 2c_i(T) \boxed{>} \hat{\mu}_i(T) + c_i(T)$$

$\boxed{>} + \boxed{>}$: $\forall j, \quad \hat{\mu}_j(T) - c_j(T) < \mu_j < \hat{\mu}_j(T) + c_j(T)$

$\boxed{>}$: $\forall i \ge 2, \quad c_i(T) < \dfrac{\mu_1 - \mu_i}{2}$

We need the following event to happen whp

$$\forall 2 \le i \le k, \quad \hat{\mu}_1(T) + c_1(T) > \hat{\mu}_i(T) + c_i(T)$$

A sufficient condition for this is

$$\hat{\mu}_1(T) + c_1(T) \boxed{>} \mu_1 \boxed{>} \mu_i + 2c_i(T) \boxed{>} \hat{\mu}_i(T) + c_i(T)$$

$$\boxed{>} + \boxed{>} : \forall j, \quad \hat{\mu}_j(T) - c_j(T) < \mu_j < \hat{\mu}_j(T) + c_j(T)$$

$$\boxed{>}: \forall i \ge 2, \quad c_i(T) < \frac{\mu_1 - \mu_i}{2}$$

Trade-off on $c_j(T)$

We apply Hoeffding inequality to bound the probability of

We apply Hoeffding inequality to bound the probability of

$$\forall j, \forall t \leq T, \quad \hat{\mu}_j(t) - c_j(t) < \mu_j < \hat{\mu}_j(t) + c_j(t)$$

We apply Hoeffding inequality to bound the probability of

$$\forall j, \forall t \leq T, \quad \hat{\mu}_j(t) - c_j(t) < \mu_j < \hat{\mu}_j(t) + c_j(t)$$

We apply Hoeffding inequality to bound the probability of

$$\forall j, \forall t \leq T, \quad \hat{\mu}_j(t) - c_j(t) < \mu_j < \hat{\mu}_j(t) + c_j(t)$$

$$\Pr[\,|\hat{\mu}_j(t) - \mu_j| > c_j(t)] \leq 2\exp\left(-\frac{2c_j^2}{n_j(1/n_j)^2}\right) = 2\exp(-2c_j^2 n_j)$$

We apply Hoeffding inequality to bound the probability of

$$\forall j, \forall t \leq T, \quad \hat{\mu}_j(t) - c_j(t) < \mu_j < \hat{\mu}_j(t) + c_j(t)$$

$$\Pr[\,|\hat{\mu}_j(t) - \mu_j| > c_j(t)] \leq 2\exp\left(-\frac{2c_j^2}{n_j(1/n_j)^2}\right) = 2\exp(-2c_j^2 n_j)$$

So the Hoeffding inequality suggests us to choose

$$c_j(T) = \Omega(\sqrt{\frac{\log T}{n_j(T)}})$$

For this choice of $c_i(T)$, the condition $c_i(T) < \dfrac{\mu_1 - \mu_i}{2}$ becomes to

For this choice of $c_i(T)$, the condition $c_i(T) < \frac{\mu_1 - \mu_i}{2}$ becomes to

$$\sqrt{\frac{c \log T}{n_i(T)}} < \frac{\mu_1 - \mu_i}{2}$$

For this choice of $c_i(T)$, the condition $c_i(T) < \frac{\mu_1 - \mu_i}{2}$ becomes to

$$\sqrt{\frac{c \log T}{n_i(T)}} < \frac{\mu_1 - \mu_i}{2}$$

This means $n_i(T) = \Omega(\log T)$, so we need to try each arm $\Omega(\log T)$ times for free!

For this choice of $c_i(T)$, the condition $c_i(T) < \frac{\mu_1 - \mu_i}{2}$ becomes to

$$\sqrt{\frac{c \log T}{n_i(T)}} < \frac{\mu_1 - \mu_i}{2}$$

This means $n_i(T) = \Omega(\log T)$, so we need to try each arm $\Omega(\log T)$ times for free!

Some tedious calculations are required to obtain the final regret bound, which is $\Theta(\log T)$