Lec 14

Friday, October 3, 2025 5:57 AM

* HW-4 due Oct 16

* Recap: LPs, Approx alg (using LPs)

Randomized Algorithms

Algorithm has access to random bits ← assume 1 unit operation to generate Bernoulli$(p) = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{o.w.} \end{cases}$

Basic Tools:

① **Linearity of Expectation**: For any two random vars X, Y:

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$ ← even if correlated

$\mathbb{E}[XY]$ need not always be $\mathbb{E}X \cdot \mathbb{E}Y$ ← for indep r.v.s this is equality

E.g., we can write running time of rand alg as
Time $= X_1 + X_2 + \ldots$ and take expectations.

② **Union Bound**: For any two events A and B depending on r.v. X,

$$P[A \cup B] \leq P[A] + P[B]$$

E.g., suppose your alg calls two functions, each of which fails with probability $\leq \epsilon$
$\Rightarrow$ your alg fails w.p. $\leq 2\epsilon$

③ **Markov's Inequality**: For any non-negative r.v. X,

$$P[X \geq k \cdot \mathbb{E}[X]] \leq \frac{1}{k} \quad \text{for } k \geq 1$$

Pf: $\mathbb{E}[X] = \underbrace{P[X \geq \alpha] \cdot \mathbb{E}[X \mid X \geq \alpha]}_{\geq P[X > \alpha] \cdot \alpha} + \underbrace{P[X < \alpha] \cdot \mathbb{E}[X \mid X < \alpha]}_{\geq 0}$ ∎

E.g., if $\mathbb{E}[\text{Running time}] = 10n^2$ then
$$P[\text{Running time} > 20n^2] \leq \frac{1}{2}$$

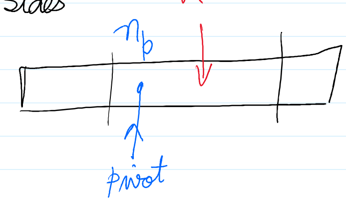
Finding Median

Given n numbers $A[0, \ldots n-1]$. ← Say distinct for simplicity

Given $k \in \{1, 2, \ldots, n\}$

Goal: Find the k-th largest number ← $(k-1)$ smaller #s & $n-k$ large #s

Idea: Find a $\underbrace{\text{good pivot}}_{\geq 30\% \text{ elems larger and smaller}}$ and discard one of its sides



Previously we saw Median-of-Median to find a good pivot recursively. ⟵ complicated

Idea: Just choose one of n numbers uniformly at random as the next pivot. ⟵ 40% chance of being a good pivot

Takes $\log n$ time to generate

What is the running time of this algorithm?

Time is a r.v.

(A) What is its expectation?

(B) How likely is the time to be close to expectation.

(A) Thm: $\mathbb{E}[\text{Time of alg}] = O(n)$.

Plan: Write Time = Sum of 'simple' r.v.s X_i and use linearity of expectation.

Note: Input size $n \longrightarrow \frac{7}{10} n$ on finding a good pivot

Also, suppose the alg takes cn time in an iteration (i.e. in trying a pivot)

$$\text{Time} \leq cn \cdot X_1 + c \cdot \left(\frac{7}{10}\right) n X_2 + c \left(\frac{7}{10}\right)^2 n X_3 + c \cdot \left(\frac{7}{10}\right)^3 n X_4 + \cdots$$

where r.v. X_i denotes number of attempts between $(i-1)$ and i-th good pivots.

$p = 0.4$

Obs: $\mathbb{E}[X_i] = 1/0.4 = 5/2$ because $\mathbb{E}[X_i] = p \cdot 1 + (1-p)p \cdot 2 + (1-p)^2 \cdot p \cdot 3 + \cdots = 1/p$

Obs: $\mathbb{E}[X_i] = 1/0.4 = 5/2$ because $\mathbb{E}[X_i] = p.1 + (1-p)p.2 + (1-p)^2 \cdot p \cdot 3 + \dots = 1/p$ $p=0.4$

$$\Rightarrow \mathbb{E}[\text{Time}] \le Cn \cdot \frac{5}{2}\left(1 + \frac{7}{10} + \left(\frac{7}{10}\right)^2 + \left(\frac{7}{10}\right)^3 + \dots\right)$$

$$\le Cn \cdot \frac{5}{2} \cdot \frac{1}{1-7/10} = O(n) \quad \blacksquare$$

Takeaway: ① Write runtime as sum of simple r.v.s

② Use linearity of expectation

Ⓑ How to argue this randomized alg works with "high" probability?

We showed $\mathbb{E}[\text{Time}] \le Cn$ for some (different) constant C

Q1: What is $P[\text{Time} \ge 2Cn]$? At most $1/2$ by Markov's ineq

Q2: What is $P[\text{Time} \ge k(2Cn)]$? By Markov's ineq $\frac{1}{2k}$ (like $k=100$)

Can we do better?

New Alg: Run old alg from scratch for $2\boxed{Cn}$ steps. ($\mathbb{E}[\text{Time}]$)
Repeat if no answer found.

New Alg succeeds if either of the k-rand old alg succeeds

$$\Rightarrow P[\text{New Alg takes time} \ge k \cdot 2Cn] \le \frac{1}{2^k} \quad \text{for any positive integer } k$$

Takeaway: We can make failure probability decay very fast. $P[\text{Time} \ge k\, 2\,\mathbb{E}[\text{Time}]] \le \frac{1}{2^k}$.