CSE319: Modern Algorithm Design

Homework 1 Solutions

Submitted By: Divyajeet Singh (2021529) **Discussion Partner:** Siddhant Rai Viksit (2021565)

Solution 1.

The given algorithm can be described by Algorithm (1).

Algorithm 1 Modified Quicksort Algorithm

---

procedure Mod-Quicksort($A[l:r]$):
    if $l \geq r$ **then**
        return
    end if
    while True **do**
        $i \sim l, \ldots, r$ uniformly at random
        $j \leftarrow$ Partition($A[l:r], i$)
        if $j - l + 1 \leq \frac{1}{4}(r - l + 1)$ **and** $r - j + 1 \leq \frac{3}{4}(r - l + 1)$ **then**
            break
        end if
    end while
    Mod-Quicksort($A[l:j-1]$)
    Mod-Quicksort($A[j+1:r]$)
end procedure

Part (a)

To analyze the expected running time of this algorithm, we first bound the probability of successfully finding a *good pivot*[1]. Let's find the number of good pivots in an array of size n. It is easy to see that if the sorted index j of the pivot is less than $\frac{1}{4}n$, then more than $\frac{3}{4}n$ elements land to its right. Similarly, if $j > \frac{3}{4}n$, then more than $\frac{3}{4}n$ elements land to its left. So, we see that the number of good pivots in an array of size n is given by

$$\left(\frac{3}{4} - \frac{1}{4}\right) n = \frac{n}{2} \tag{1}$$

Since the pivot is picked uniformly at random, the probability of picking a good pivot in one trial is

$$p = \frac{n}{2n} = \frac{1}{2} \tag{2}$$

It is not hard to notice that the number of trials required for successfully picking a pivot is distributed as a random variable $X \sim$ Geometric(p) since we repeat the trials (which are independent) until a good pivot is found. We find the expected number of trials required to get a successful trial, i.e. find a good pivot.

$$\mathbb{E}[X] = p^{-1} = \left(\frac{1}{2}\right)^{-1} = 2 \tag{3}$$

Note that each time a pivot is picked and tested for *goodness*, it takes $O(n)$ time for an array of size n since its rank is computed using partitioning (or any other $O(n)$ time algorithm). Finally, we estimate the rank of a good pivot.

[1] **Definition.** ***Good Pivot*:** A pivot having at most $\frac{1}{4}$ of the total elements to one side.

We calculate the number of elements R that can land to the right (or left) of a good pivot. Let $\mathbf{G}$ be the event that a good pivot is chosen. Then, $\mathbf{G} = \left\{\frac{1}{4}n \leq R \leq \frac{3}{4}n\right\}$. Since the pivot is chosen uniformly at random, we have

$$\mathbb{P}[R = i \mid \mathbf{G}] = \begin{cases} \frac{2}{n} & \text{if } \frac{1}{4}n \leq i \leq \frac{3}{4}n \\ 0 & \text{otherwise} \end{cases} \qquad \text{by (1)} \tag{4}$$

$$\mathbb{E}[R \mid \mathbf{G}] = \sum_{i=1}^{n} i\, \mathbb{P}[R = i \mid \mathbf{G}] = \sum_{i=\frac{1}{4}n}^{\frac{3}{4}n} \frac{2i}{n} = \frac{2}{n} \sum_{i=\frac{1}{4}n}^{\frac{3}{4}n} i = \frac{2}{n}\left[\frac{1}{2} \cdot \left(\frac{n}{2} + 1\right) \cdot \left(2 \cdot \frac{n}{4} + \frac{n}{2}\right)\right] = \frac{n}{2} + 1 \approx \frac{n}{2} \tag{5}$$

Similarly, approximately $\frac{n}{2}$ elements land to the left of a good pivot, which means that the expected rank of a good pivot is $\frac{n}{2}$. Let $T(n)$ denote the time taken by Algorithm (1) on an input of size n. The recurrence relation can be defined as follows (the second term represents the expected work done to find the pivot).

$$\begin{aligned}
\mathbb{E}[T(n)] &= 2\mathbb{E}_{k\sim\frac{1}{4}n,\ldots,\frac{3}{4}n}[T(k)] + 2cn = 2\mathbb{E}\left[T\left(\frac{n}{2}\right)\right] + 2cn \\
&= 2\left(2\mathbb{E}_{k\sim\frac{1}{4}\cdot\frac{n}{2},\ldots,\frac{3}{4}\cdot\frac{n}{2}}[T(k)] + 2c\cdot\frac{n}{2}\right) + 2cn = 2\left(2\mathbb{E}\left[T\left(\frac{n}{4}\right)\right] + 2c\cdot\frac{n}{2}\right) + 2cn \\
&= 2\left\{2\left(2\mathbb{E}\left[T\left(\frac{n}{8}\right)\right] + 2c\cdot\frac{n}{4}\right) + 2c\cdot\frac{n}{2}\right\} + 2cn \\
&= 2^3\ \mathbb{E}\left[T\left(\frac{n}{2^3}\right)\right] + 2^2\cdot 2c\cdot\frac{n}{2^2} + 2^1\cdot 2c\cdot\frac{n}{2^1} + 2^0\cdot 2c\cdot\frac{n}{2^0} \\
&\ \vdots \\
&\leq 2^k\ \mathbb{E}\left[T\left(\frac{n}{2^k}\right)\right] + (k+1)\cdot 2cn
\end{aligned} \tag{6}$$

The recurrence ends when

$$\frac{n}{2^k} = 1 \implies 2^k = n \implies k = \log_2 n \tag{7}$$

Substituting (7) in (6) gives the expected time complexity of the modified quicksort algorithm where we only partition around good pivots.

$$\begin{aligned}
\mathbb{E}[T(n)] &\leq 2^{\log_2 n}\mathbb{E}[T(1)] + 2cn\cdot(\log_2 n + 1) \\
&= 2kcn\log_2 n + (2c+1)\cdot n \qquad \text{assuming } T(1) = 1 \\
&\leq c^2\cdot n\log_2 n \quad \forall\ n \geq 2 \\
\therefore \mathbb{E}[T(n)] &= O(n\log n)
\end{aligned} \tag{8}$$

Hence, the modified quicksort algorithm runs in $O(n\log n)$ time in expectation.

Part (b)

The bulk of the algorithm runtime comes from calulating the ranks of the randomly chosen pivots to identify whether they are good, and the partitioning steps. Let $T_i^{(j)} \sim \textsc{Geometric}\left(\frac{1}{2}\right)$ denote the number of trials to find a good pivot at the j^{th} branch of the i^{th} level/depth of recursion. There will be $O(\log n)$ levels of recursion, since the size of the array goes down by at least $\frac{3}{4}$ at each level. Then, the total work done by the algorithm is given by

$$W = \sum_{i=1}^{k\log n}\sum_{j=1}^{2^{i-1}} c\cdot n_i^{(j)} T_i^{(j)} \tag{9}$$

for some constants $c, k > 0$, where $n_i^{(j)}$ denotes the size of the sub-array at the j^{th} branch of the i^{th} level of recursion. For a sanity check, we calculate the expectation of W to verify that it is $O(n\log n)$.

$$\mathbb{E}[W] = c\sum_{i=1}^{k\log n}\sum_{j=1}^{2^{i-1}}\mathbb{E}\left[n_i^{(j)}\right]\cdot\mathbb{E}\left[T_i^{(j)}\right] = c\sum_{i=1}^{k\log n}\sum_{j=1}^{2^{i-1}}\frac{n}{2^i}\cdot 2 = c\cdot n\sum_{i=1}^{k\log n}\sum_{j=1}^{2^{i-1}}\frac{1}{2^{i-1}} = c\cdot n\sum_{i=1}^{k\log n} 1 = ck\cdot n\log n \tag{10}$$

where $n_i^{(j)}$ depends on the expected rank of the pivot (using (5)) and we use the independence of the size of the sub-array at any step and the number of trials used to identify a good pivot. So, the expected work done by the algorithm is $O(n\log n)$.

We bound the probability that (ignoring constants for simplicity)

$$\begin{aligned}
\mathbb{P}\left[\sum_{i=1}^{\log n}\sum_{j=1}^{2^{i-1}} n_i^{(j)} T_i^{(j)} > n\log n\right] &\leq \mathbb{P}\left[\exists i : \sum_{j=1}^{2^{i-1}} T_i^{(j)} > 2\log n\right] \\
&\leq \sum_{i=1}^{\log n} \mathbb{P}\left[\sum_{j=1}^{2^{i-1}} T_i^{(j)} > 2\log n\right] \qquad \text{by union bound} \\
&\leq \sum_{i=1}^{\log n} \mathbb{P}\left[\exists j : T_i^{(j)} > 2\log n\right] \\
&\leq \sum_{i=1}^{\log n}\sum_{j=1}^{2^{i-1}} \mathbb{P}\left[T_i^{(j)} > 2\log n\right] \qquad \text{by union bound} \\
&= \sum_{i=1}^{\log n} 2^{i-1} \cdot \left(\frac{1}{2}\right)^{2\log n} \\
&= \frac{1}{n^2} \cdot (2^{\log n} - 1) \approx \frac{n}{n^2} = \frac{1}{n}
\end{aligned} \tag{11}$$

where the first inequality holds as the probability that the total work done exceeds $n \log n$ is at most the probability that we use more than $O(\log n)$ trials at some level to find a good pivot. Therefore,

$$\mathbb{P}[W \leq n\log n] = 1 - \mathbb{P}[W > n\log n] \geq 1 - \frac{1}{n} \tag{12}$$

Solution 2.

Part (a)

Given two vectors $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_d)$ chosen independently and uniformly at random from $\{-1, 1\}^d$. We can derive the required result using a Chernoff bound on the inner product. For each $1 \leq i \leq d$, x_i and y_i are uniform random variables taking values $\{-1, 1\}$. Let $Z_i = x_i y_i$ for each i. Then the random variable Z_i denotes the i^{th} component of the inner product $\langle \mathbf{x}, \mathbf{y} \rangle$. Thus, the inner product can be written as

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{d} x_i y_i = \sum_{i=1}^{d} Z_i = Z \tag{13}$$

where the random variable Z, the sum of d independent random variables Z_i, is the inner product of $\mathbf{x}$ and $\mathbf{y}$. We first note that

$$\mathbb{P}[|Z| \geq \epsilon d] = \mathbb{P}[Z \geq \epsilon d] + \mathbb{P}[Z \leq -\epsilon d] = 2\mathbb{P}[Z \geq \epsilon d] \tag{14}$$

where the last equality holds by symmetry[2]. Each Z_i is a uniform random variable taking values $\{-1, 1\}$, as

$$\begin{aligned}
\mathbb{P}[Z_i = 1] = \mathbb{P}[x_i = 1, y_i = 1] + \mathbb{P}[x_i = -1, y_i = -1] = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2} \\
\mathbb{P}[Z_i = -1] = \mathbb{P}[x_i = 1, y_i = -1] + \mathbb{P}[x_i = -1, y_i = 1] = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2}
\end{aligned} \tag{15}$$

where we crucially use the independence of x_i and y_i. To derive a Chernoff bound, we find the moment generating function of Z_i $(1 \leq i \leq d)$ and Z.

$$\Phi_{Z_i}(t) = \mathbb{E}[e^{tZ_i}] = \mathbb{P}[Z_i = 1]e^t + \mathbb{P}[Z_i = -1]e^{-t} = \frac{e^t + e^{-t}}{2} = \cosh t \tag{16}$$

$$\Phi_Z(t) = \prod_{i=1}^{d} \Phi_{Z_i}(t) = \prod_{i=1}^{d} \cosh t = (\cosh t)^d \leq \left(e^{\frac{t^2}{2}}\right)^d = e^{\frac{t^2 d}{2}} \qquad \left(\cosh x \leq e^{\frac{x^2}{2}} \ \forall\, x \in \mathbb{R}\right) \tag{17}$$

using the independence of Z_i for all $(1 \leq i \leq d)$.

[2] See **Appendix** for the proof of symmetry.

We now bound the probability of the inner product exceeding ϵd.

$$\begin{aligned}\mathbb{P}[Z \geq \epsilon d] = \mathbb{P}[e^{tZ} \geq e^{t\epsilon d}] &\leq \frac{\Phi_Z(t)}{e^{t\epsilon d}} \quad (t > 0) \\ &\leq e^{-t\epsilon d} \cdot e^{\frac{t^2 d}{2}} = e^{\frac{t^2 d}{2} - t\epsilon d} \\ &\leq \min_{t>0} \exp\left(\frac{t^2 d}{2} - t\epsilon d\right)\end{aligned} \tag{18}$$

where the first inequality follows by Markov's inequality. To find the minimum, we differentiate the exponent with respect to t and set it to zero.

$$\frac{d}{dt}\left(\frac{t^2 d}{2} - t\epsilon d\right) = td - \epsilon d = 0 \implies t = \epsilon \tag{19}$$

Substituting $t = \epsilon$ gives

$$\mathbb{P}[Z \geq \epsilon d] \leq \exp\left(\frac{\epsilon^2 d}{2} - \epsilon^2 d\right) = \exp\left(-\frac{\epsilon^2 d}{2}\right) \leq \exp\left(-\frac{\epsilon^2 d}{6}\right) \tag{20}$$

where the last inequality follows since e^{-x} is monotonically decreasing. Thus, by (14) and (20), we have

$$\mathbb{P}[|\langle \mathbf{x}, \mathbf{y}\rangle| \geq \epsilon d] \leq 2\exp\left(-\frac{\epsilon^2 d}{6}\right) \tag{21}$$

Part (b)

For a constant $\epsilon > 0$, a set S of unit vectors is called ϵ-orthonormal if[3] for any $\hat{\mathbf{x}} \neq \hat{\mathbf{y}} \in S$, $|\langle \hat{\mathbf{x}}, \hat{\mathbf{y}}\rangle| \leq \epsilon$. For a constant c and $d_0 \geq 0$, we sample $N = e^{c\epsilon^2 d}$ unit vectors from $\left\{-\frac{1}{\sqrt{d}}, +\frac{1}{\sqrt{d}}\right\}^d$ independently and uniformly. Let

$$\hat{\mathbf{x}} = \frac{1}{\sqrt{d}}\mathbf{x} = \frac{1}{\sqrt{d}}(x_1, x_2, \ldots, x_d) \qquad \hat{\mathbf{y}} = \frac{1}{\sqrt{d}}\mathbf{y} = \frac{1}{\sqrt{d}}(y_1, y_2, \ldots, y_d) \tag{22}$$

denote any two unit vectors in S, using the same notation as **Solution 2 Part (a)**. The inner product of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ is given by

$$\langle \hat{\mathbf{x}}, \hat{\mathbf{y}}\rangle = \sum_{i=1}^{d} \frac{x_i}{\sqrt{d}} \cdot \frac{y_i}{\sqrt{d}} = \frac{1}{d}\sum_{i=1}^{d} Z_i = \frac{Z}{d} \tag{23}$$

By (21) (**Solution 2 Part (a)**), we see (since $d > 0$)

$$\mathbb{P}[|Z| \geq \epsilon d] = \mathbb{P}\left[\left|\frac{Z}{d}\right| \geq \epsilon\right] \leq 2\exp\left(-\frac{\epsilon^2 d}{6}\right) \tag{24}$$

Let us bound the probability that there is some pair of vectors in S that is not ϵ-orthogonal.

$$\begin{aligned}\mathbb{P}[\exists \hat{\mathbf{x}}, \hat{\mathbf{y}} \in S : |\langle \hat{\mathbf{x}}, \hat{\mathbf{y}}\rangle| \geq \epsilon] &\leq \binom{N}{2}\, \mathbb{P}[|\langle \hat{\mathbf{x}}, \hat{\mathbf{y}}\rangle| \geq \epsilon] \quad \text{by union bound} \\ &= \frac{N(N-1)}{2} \cdot 2\exp\left(-\frac{\epsilon^2 d}{6}\right) \\ &\leq N^2 \exp\left(-\frac{\epsilon^2 d}{6}\right) = \exp\left(2c\epsilon^2 d - \frac{\epsilon^2 d}{6}\right)\end{aligned} \tag{25}$$

We want that such a pair with a large inner product exists with probability at most half. So,

$$\begin{aligned}\exp\left(2c\epsilon^2 d - \frac{\epsilon^2 d}{6}\right) &\leq \frac{1}{2} \\ 2c\epsilon^2 d - \frac{\epsilon^2 d}{6} &\leq -\ln 2 \\ c &\leq \frac{1}{12} - \frac{1}{2\epsilon^2 d}\ln 2\end{aligned} \tag{26}$$

[3]We have slightly modified the definition from the problem, since the pairs $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ cannot be ϵ-orthogonal for any $\epsilon > 1$ otherwise, as their inner product is always 1.

Since we sample $N = e^{c\epsilon^2 d}$ vectors, $c \geq 0$. Thus, we want

$$\frac{1}{12} \gg \frac{1}{2\epsilon^2 d} \ln 2 \implies d \gg \frac{6}{\epsilon^2} \ln 2 \tag{27}$$

Since we just want to prove the existence of such constants c and d_0, we can choose

$$d \geq d_0 = \frac{60}{\epsilon^2} \ln 2 \implies c \leq \frac{1}{12} - \frac{1}{120} = \frac{3}{40} \tag{28}$$

Therefore, for $d_0 = \frac{60}{\epsilon^2} \ln 2$, we can sample $N = e^{\frac{3}{40}\epsilon^2 d}$ unit vectors in $d \geq d_0$ dimensions such that with probability at least half, the set of vectors is ϵ-orthonormal.

Appendix

Solution 2 (a)

The derivation for the bound on the lower tail of the inner product is given here.

$$\begin{aligned}
\mathbb{P}[Z \leq -\epsilon d] = \mathbb{P}[e^{tZ} \geq e^{-t\epsilon d}] &\leq \frac{\Phi_Z(t)}{e^{-t\epsilon d}} \quad (t < 0) \\
&\leq e^{t\epsilon d} \cdot e^{\frac{t^2 d}{2}} = e^{\frac{t^2 d}{2} + t\epsilon d} \\
&\leq \min_{t<0} \exp\left(\frac{t^2 d}{2} + t\epsilon d\right)
\end{aligned} \tag{29}$$

We differentiate with respect to t and set it to zero to find the minimum.

$$\frac{d}{dt}\left(\frac{t^2 d}{2} + t\epsilon d\right) = td + \epsilon d = 0 \implies t = -\epsilon \tag{30}$$

Substituting $t = -\epsilon$ gives

$$\mathbb{P}[Z \leq -\epsilon d] \leq \exp\left(\frac{\epsilon^2 d}{2} - \epsilon^2 d\right) = \exp\left(-\frac{\epsilon^2 d}{2}\right) \leq \exp\left(-\frac{\epsilon^2 d}{6}\right) \tag{31}$$

which is the same as the derived bound of the upper tail.