6.856 — Randomized Algorithms

David Karger

Handout #8, September 30, 2002 — Homework 3 Solutions

Problem 1

We may think of asking a resident as flipping a coin with bias p=f. Flip the coin N times. If you get k heads, set $\hat{p} = k/N$. Note k has a binomial distribution with mean $\mu = pN$. Thus, using Chernoff bounds:

$$\begin{aligned}
\Pr[|p - \hat{p}| > \epsilon p] &= \Pr[|\mu - k| > \epsilon\mu] \\
&= \Pr[\mu - k > \epsilon\mu] + \Pr[k - \mu > \epsilon\mu] \\
&= \Pr[k < (1-\epsilon)\mu] + \Pr[k > (1+\epsilon)\mu] \\
&\leq e^{-\epsilon^2\mu/2} + e^{-\epsilon^2\mu/3} \\
&\leq 2e^{-\epsilon^2\mu/3} \\
&= 2e^{-\epsilon^2 Np/3}
\end{aligned}$$

(the constant 3 in the second to last line follows from the fact that we are assuming $\epsilon < 1$). Set this bound equal to δ, and solve for N to find that

$$N = 3\ln(2/\delta)/(\epsilon^2 p)$$

trials suffice. Since $p \geq a$ by assumption, certainly

$$N = 3\ln(2/\delta)/(\epsilon^2 a)$$

trials suffice.

An interesting additional exercise is to show that even if we do not have any a priori bound a on p, we can still estimate p as above in $O(\ln(2/\delta)/(\epsilon^2 p))$ trials with probability exceeding $1 - \delta$.

Problem 2

We show that with high probability, every element is compared to $O(\log n)$ pivots. This proves that there are $O(n \log n)$ comparisons overall. To prove our claim, take a particular element x and consider the series of recursively defined subproblems $S = S_1, S_2, \ldots$ into which element x is placed. Subproblem S_{k+1} is constructed from S_k by choosing a random pivot element from S_k and placing into S_{k+1} all elements on the same side of the pivot as x.

Of course, once S_k has size 1, it contains only 1 element, namely x, and the recursion stops. We show that with high probability, there is an $r = O(\log n)$ such that S_r has exactly one element. This proves our initial claim. To do so, call a subproblem S_k *good* if $|S_{k+1}| \leq \frac{3}{4}|S_k|$. Since $|S_1| = n$, if the sequence $S_1, \ldots, S_r$ contains $\log_{4/3} n$ good subproblems then $|S_r| \leq 1$. So we just need to show that with high probability $S_1, \ldots, S_r$ has $\log_{4/3} n$ good subproblems.

To do this, note that since the pivot that yields S_{k+1} is chosen uniformly at random from S_k, the probability that S_k is good is at least $1/2$, independent of the goodness of all the other subproblems (a slight subtlety here: conditioning on the goodness of previous problems can bias the probability of S_{k+1} being good, but cannot bring the probability below $1/2$). It follows that the number of good subproblems in the sequence $S_1, \ldots, S_r$ stochastically dominate a binomial distribution with mean at least $r/2$. Thus, the probability of fewer than, say, $r/4$ good subproblems is, by the Chernoff bound, at most $e^{-(1/2)^2(r/2)/2} = e^{-r/16}$. It follows that if we take $r = 32 \ln n$, then since $\log_{4/3} n < 16 \ln n$,

$$
\begin{aligned}
\Pr[|S_r| > 1] &\leq \Pr[\text{less than } \log_{4/3} n \text{ good subproblems}] \\
&\leq \Pr[\text{less than } 16 \ln n \text{ good subproblems}] \\
&\leq 1/n^2
\end{aligned}
$$

Thus the probability that any one of the n elements encounters more than $32 \log n$ pivots is less than $1/n$.

A common mistake was to assume that the the variables X_{ij}, defined to be 1 if j is a pivot to which i compares, were independent. The only approach on this problem set that lead to a solid alternate solution on was to show that $\{X_{ij} | i < j\}$ and $\{X_{ij} | i > j\}$ are independent, and apply Chernoff to them separately.

Problem 3

(a) Suppose we throw k balls into n bins. Let $X_1, X_2, \ldots, X_k$ be random variables, so that X_i is 1 if the i-th ball lands in a bin by itself, and 0 otherwise. The probability that a certain ball lands in a bin by itself is equal to the probability that all other balls land in different bins, i.e. $(1 - 1/n)^{k-1}$.

The total number of balls that land in a bin by themselves are therefore $\sum X_i$, and we have to determine $E[\sum X_i]$. By linearity of expectation, we have

$$E\left[\sum X_i\right] = kE[X_1] = k\left(1 - \frac{1}{n}\right)^{k-1}$$

So if $k = \varepsilon n$, then the number of bins with only one ball in them is

$$n\varepsilon - E\left[\sum X_i\right] = n\varepsilon(1 - 1/n)^{n\varepsilon - 1} \geq n\varepsilon / e^{\varepsilon}.$$

Thus, the number of balls we are expected to keep is at most

$$n\varepsilon - n\varepsilon/e^{\varepsilon} = n\varepsilon(1 - 1/e^{\varepsilon}).$$

(b) If everything went according to expectation, then after i rounds the size of our set would be $n\varepsilon_i$, where ε_i satisfies the recurrence

$$\varepsilon_{i+1} = \varepsilon_i(1 - \exp(-\varepsilon_i)) \leq \varepsilon_i\,(1 - (1 - \varepsilon_i)) = \varepsilon_i^2$$

This implies that

$$\log \varepsilon_{i+1} \leq 2 \log \varepsilon_i. \tag{1}$$

Obviously, we have $\varepsilon_0 = 1$, but the result from (a) shows that after one round, we expect $\varepsilon_1 = 1 - 1/e \approx 0.63212$. With (1) this implies

$$\log \varepsilon_k \leq 2^{k-1} \log(1 - 1/e) \tag{2}$$

The number of rounds until we have no elements left is equal to the smallest k such that $n\varepsilon_k < 1$, which is equivalent to $\log \varepsilon_k \leq -\log n$. Using (2) this works out to $\log(1 - 1/e) \cdot 2^{k-1} \leq -\log n$, and taking log's yields $k = \Theta(\log \log n)$.

(c) In expectation, $< 1 - e^{-1}$ remain after 1 round. So, by the Markov Inequality,

$$Pr[\text{remaining fraction} > c\mu] < \frac{1}{c}, c > 1$$

$$\Rightarrow \text{ after } O(\lg \lg n) \text{ rounds, } Pr[\text{remaining fraction} > c'(1 - e^{-1})] < (\frac{1}{c})^{O(\lg \lg n)}$$

for appropriately chosen c'. So w.h.p (1-o(1)), we reach $\epsilon = c'(1 - e^{-1})$ after $O(\lg \lg n)$ rounds.

Suppose that instead of going from ϵ to ϵ^2, we were only able to reduce to $t\epsilon^2$ for constant $t > 1$. Setting up the recurrence and solving it as in (b), it can be shown that for appropriately chosen t, we still require $O(\lg \lg n)$ rounds to complete if everything proceeds according to expectation. Define a good round to be one in which the fraction remaining is less than t times the expected. Then, by the Markov Inequality, we bound the likelihood of a bad round:

$$Pr[\text{fraction remaining} > t\mu] < \frac{1}{t} \Rightarrow Pr[\text{good round}] > 1 - 1/t$$

Hence, we require $< \frac{1}{1-1/t} O(\lg \lg n) = O(\lg \lg n)$ rounds in expectation to get rid of all balls. Since the rounds are independent, and we can think of "goodness" as an indicator variable, we may Chernoff bound the number of rounds required:

$$Pr[\text{rounds} > d \lg \lg n] < 2^{-D \lg \lg n} < 2^{-D \lg \lg n} = \frac{1}{(\lg n)^D}$$

for some constants d and D.

Thus w.h.p(1-o(1)), it takes another $O(\lg\lg n)$ rounds from $\epsilon = c'(1-e^{-1})$ to get rid of all balls. Therefore, overall, the process takes $O(\lg\lg n)$ rounds w.h.p, since $(1-o(1))(1-o(1)) = 1-o(1)$.

A common mistake was to define a good round to be one in which we reach the expected fraction for the i-th round on the i-th round. These rounds are not independent, since previous rounds clearly affect how likely it is to reach the expected fraction for the current round. Many also forgot to show that it takes $O(\log\log n)$ steps to get from 1 to a constant fraction.

Problem 5

(a) Consider the function

$$g(\lambda) = \lambda e^{tx} + (1-\lambda)e^{ty} - e^{t(\lambda x+(1-\lambda)y)}.$$

We show that for any fixed t, x and $y > x$, $g(\lambda) > 0$ in the interval $0 < \lambda < 1$. This proves the claim. First note that $g(0) = g(1) = 0$. We now prove that there are no other $\hat{\lambda} \in [0,1]$ for which $g(\hat{\lambda}) = 0$. Suppose there were. Then by Rolle's theorem, there would have to be two zeroes of $g'(\lambda)$, one in the interval $[0,\hat{\lambda}]$ and one in the interval $[\hat{\lambda},1]$. We show in contradiction that there is at most one such point. To see this, note that

$$g'(\lambda) = e^{tx} - e^{ty} - e^{t(\lambda x+(1-\lambda)y)}t(x-y).$$

Solving, we find that $g'(\lambda) = 0$ only if

$$e^{t(\lambda x+(1-\lambda y))} = \frac{e^{ty}-e^{tz}}{t(y-x)}.$$

It is easy to verify that the left hand side is monotonic in λ, meaning the equation has a unique solution in λ.

(b) Observe that if we replace λ the previous step by our random variable Z, since $0 \le Z \le 1$,

$$\begin{aligned} f(Z) &= f((1-Z)\cdot 0 + Z\cdot 1) \\ &\le (1-Z)\cdot f(0) + Z\cdot f(1). \end{aligned}$$

Now for random variables X and Y, if $X \le Y$, it is easy to prove that $E[X] \le E[Y]$. It follows that

$$\begin{aligned} E[f(Z)] &\le E[(1-Z)\cdot f(0) + Z\cdot f(1)] \\ &= (1-p)\cdot f(0) + p\cdot f(1) \\ &= E[f(X)] \end{aligned}$$

(c) Let $E[Y_i] = p$, $0 \leq p \leq 1$, Let X_i be 1 with probability p and define $X = \sum X_i$; by the previous section we know that $E[e^{tY_i}] \leq E[e^{tX_i}]$ for any t. Write $\overline{\delta} = \delta/E[Y]$ and $\mu = E[Y]$. We know that

$$\begin{aligned} \Pr[Y - E[Y] > \delta] &= \Pr[Y > (1+\overline{\delta})E[Y]] \\ &\leq \frac{E[e^{tY}]}{e^{t(1+\overline{\delta})\mu}} \\ &= \frac{\prod E[e^{tY_i}]}{e^{t(1+\overline{\delta})\mu}} \\ &\leq \frac{\prod E[e^{tX_i}]}{e^{t(1+\overline{\delta})\mu}} \end{aligned}$$

where the last line follows from our convexity argument. The last line, however, is directly out of the proof of the Chernoff bound on $\Pr[X > (1+\overline{\delta})\mu]$, so we can jump directly to the end of that analysis to derive a bound on the deviation probability. Lets focus on the harder case of $\overline{\delta}$ small (that is, less than $2e-1$). Suppose first that $\mu < n/3$. Then the error probability is at most

$$\begin{aligned} e^{-\overline{\delta}^2\mu/4} &= e^{-\delta^2/4\mu} \\ &\leq e^{-3\delta^2/4n} \end{aligned}$$

To get good bounds for $\mu > n/3$, consider the following trick: let $Z_i = 1 - Y_i$ be a random variable in the interval $[0,1]$, and consider $Z = \sum Z_i = n - \sum Y_i = n - Y$. Write $\hat{\delta} = \delta/E[Z]$. The Chernoff bound tells use that

$$\begin{aligned} \Pr[Y - E[Y] > \delta] &= \Pr[E[Z] - Z > \delta] \\ &= \Pr[Z < E[Z] - \delta] \\ &= \Pr[Z < (1-\hat{\delta})E[Z]] \\ &\leq e^{-\hat{\delta}^2 E[Z]/2} \\ &= e^{-\delta^2/2E[Z]} \\ &= e^{-\delta^2/2(n-\mu)} \end{aligned}$$

which, under the assumption that $\mu > n/3$, is at most $e^{-3\delta^2/4n}$

Problem 6

We are only going to prove the first part, i.e. giving the upper tail bound. For a proof for the lower tail, see [KMPS94].

First, let us prove $E\left[X_{i_1}X_{i_2}\cdots X_{i_k}\right] \leq E\left[X_{i_1}\right]E\left[X_{i_2}\right]\cdots E\left[X_{i_k}\right]$ (for $i_1, i_2, \ldots, i_k$ distinct integers in $[1,n]$).

First we know that

$$E\left[X_{i_1}\right] = E\left[X_{i_2}\right] = \cdots = E\left[X_{i_k}\right] = \left(1 - \frac{1}{n}\right)^n$$

As well,

$$\begin{aligned} E\left[X_{i_1} X_{i_2} \cdots X_{i_k}\right] &= \Pr\left(X_{i_1} \cap X_{i_2} \cap \ldots \cap X_{i_k}\right) \\ &= \left(1 - \frac{k}{n}\right)^n \quad \text{since } i_1, i_2, \ldots, i_k \text{ are distinct} \end{aligned}$$

We now note that

$$\left(1 - \frac{1}{n}\right)^k = 1 - \frac{k}{n} + \binom{k}{2}\left(\frac{1}{n}\right)^2 - \cdots \geq 1 - \frac{k}{n}$$

Thus

$$\left[\left(1 - \frac{1}{n}\right)^n\right]^k \geq \left(1 - \frac{k}{n}\right)^n$$

which implies that

$$E\left[X_{i_1}\right] E\left[X_{i_2}\right] \cdots E\left[X_{i_k}\right] \geq E\left[X_{i_1} X_{i_2} \cdots X_{i_k}\right]$$

Since each X_{i_l} only takes on a value of either 0 or 1, we know that $X_{i_l}^m = X_{i_l}$ and then

$$E\left[X_{i_1}^{m_1} X_{i_2}^{m_2} \cdots X_{i_k}^{m_k}\right] = E\left[X_{i_1} X_{i_2} \cdots X_{i_k}\right]$$

So, using the results above,

$$E\left[X_{i_1}^{m_1}\right] E\left[X_{i_2}^{m_2}\right] \cdots E\left[X_{i_k}^{m_k}\right] \geq E\left[X_{i_1}^{m_1} X_{i_2}^{m_2} \cdots X_{i_k}^{m_k}\right]$$

We now note that $E\left[X'_{i_l}\right] = E\left[X_{i_l}\right]$ and that $E\left[X'_{i_1} X'_{i_2} \cdots X'_{i_k}\right] = E\left[X'_{i_1}\right] E\left[X'_{i_2}\right] \cdots E\left[X'_{i_k}\right]$ by which we can now conclude that

$$\begin{aligned} E\left[X'^{m_1}_{i_1} X'^{m_2}_{i_2} \cdots X'^{m_k}_{i_k}\right] &= E\left[X'_{i_1}\right] E\left[X'_{i_2}\right] \cdots E\left[X'_{i_k}\right] \\ &= E\left[X_{i_1}\right] E\left[X_{i_2}\right] \cdots E\left[X_{i_k}\right] \\ &\geq E\left[X_{i_1} X_{i_2} \cdots X_{i_k}\right] \\ &\geq E\left[X_{i_1}^{m_1} X_{i_2}^{m_2} \cdots X_{i_k}^{m_k}\right] \qquad (3) \end{aligned}$$

Since the expansion of $\left(\sum_i X_i\right)^m$ contains only terms like $X_{i_1}^{m_1} X_{i_2}^{m_2} \cdots X_{i_k}^{m_k}$, we know, from the linearity of expectation that

$$E\left[\left(\sum_i X'_i\right)^k\right] \geq E\left[\left(\sum_i X_i\right)^k\right]$$

which implies that

$$E\left[Z^k\right] \leq E\left[Y^k\right]$$

Since the Taylor series expansions of e^{tZ} and e^{tY} match up term by term with terms that look like a positive (for $t \geq 0$) constant multiplied by Z^k and X^k, again, by the linearity of expectation, we can note that

$$E\left[e^{tZ}\right] \leq E\left[e^{tY}\right]$$

So, letting $\mu = E[Z]$, we get

$$\begin{aligned}
\Pr\left[Z > (1+\delta)\,\mu\right] &= \Pr\left[e^{tZ} > e^{t(1+\delta)\mu}\right] \\
&< \frac{E\left[e^{tZ}\right]}{e^{t(1+\delta)\mu}} \\
&< \frac{E\left[e^{tY}\right]}{e^{t(1+\delta)\mu}} \\
&< \left[\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right]^{\mu} \qquad (4)
\end{aligned}$$

References

[KMPS94] A. Kamath, R. Motwani, K. Palem, P. Spirakis, Tail Bounds for Occupancy and the Satisfiability Threshold Conjecture, *Proceedings FOCS 1994*, pp. 592–603.