CS5330: Randomized Algorithms

Summary of CS5330 (first half)

So far in CS5330, we have covered many randomized algorithms and many techniques for analyzing randomized algorithms. I am going to list below some of the most important techniques we have used, and also some of the most interesting algorithms. This is not intended to be comprehensive: there are important topics that we have covered (and that are in the lecture notes) that are not included here.

Contents

1 Basic Probability

Technique	**Key example**
Condition probability (and chain rule for conditional probability)	Analysis of Karger's MinCut algorithm.
Probabilistic recurrences	Analysis of the FastCut algorithm.
Expected value / linearity of expectation	Coupon collected analysis (expected number of tries): $\mathrm{E}[\text{tries}] = O(n \log n)$. Also, many, many others (e.g., balls and bins analysis, hashing, etc.)
Conditional expectation	Analysis of QuickSort (high probability version)
Principle of deferred decisions	Stable marriage analysis.
Stochastic dominance	Stable marriage analysis.

2 Tail Bounds

Technique	**Key example**
Markov's Inequality	Too many to even try to choose one.
Chebychev's Inequality	Bounding the max load on a bin, approximate counting.
Fourth Moment Method	Linear probing analysis
Chernoff Bounds	Polling/Sampling (and many, many more: coin flipping, balls and bins, hashing, load balancing, approximate median, approximating counting, two-choices analysis, etc.)

3 Algorithms

Here is an incomplete list of topics we have covered:

Problem	**Algorithms**
Min-Cut in a graph	KargerCut, FastCut
Stable marriage	Gale-Shapley Greedy matching
Sorting	QuickSort
Search trees	Treaps, random binary search trees
Hashing	Chaining (average cost, max cost), Linear probing, Cuckoo hashing
Load balancing	Max load, average load
Sampling	Polling, approximate counting, approximate median

4 Basic probability facts

In this section, I will list some basic probability facts that we have relied on over and over again in this class. For more details regarding these facts, please see the Mitzenmacher-Upfal textbook.

Fact 1 (Union Bound) *Given a collection of events $E_1, E_2, \ldots, E_n$:*

$$Pr\left[\bigcup_{j=1}^{n} E_j\right] \leq \sum_{j=1}^{n} Pr[E_j]$$

If all the events E_j are distinct, then:

$$Pr\left[\bigcup_{j=1}^{n} E_j\right] = \sum_{j=1}^{n} Pr[E_j]$$

Fact 2 (Linearity of Expectation) *Let $X_1, X_2, \ldots, X_n$ be a collection of arbitrary random variables. Then:*

$$\mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{E}[X_i] \ .$$

Fact 3 (Definition of Expectation 1) *For any random variable X that takes value in domain D:*

$$\mathrm{E}[X] = \sum_{v \in D} v \cdot Pr[X = v]$$

Fact 4 (Definition of Expectation 2) *For any non-negative random variable X that takes integer values:*

$$\mathrm{E}[X] = \sum_{j=1}^{\infty} Pr[X \geq j]$$

Fact 5 (Definition of Variance) *For any random variable X:*

$$\mathrm{Var}\, X = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$$

Fact 6 (Conditional Probability) *For any two events X and Y: $Pr[X \text{ and } Y] = Pr[X \mid Y] Pr[Y]$. For any collection of events $X_1, X_2, \ldots, X_n$:*

$$Pr\left[\bigcap_{j \in [1,n]} X_j\right] = Pr[X_1] Pr[X_2 | X_1] Pr[X_3 | X_1, X_2] Pr[X_4 | X_1, X_2, X_3] \cdots$$

Two events X and Y are ***independent*** *if $Pr[X \text{ and } Y] = Pr X Pr[Y]$.*

Fact 7 (More Variance) *Given a set of independent random variables $X_1, \ldots, X_n$:*

$$\operatorname{Var} \sum_{j=1}^{n} X_j = \sum_{j=1}^{n} \operatorname{Var} X_j$$

Given a constant a:

$$\operatorname{Var} aX = a^2 \operatorname{Var} X$$

Fact 8 (Law of Total Probability) *Given any collection of disjoint possible outcomes $B_1, X_2, \ldots B_n$ where $\sum_{j=1}^{n} Pr[B_j] = 1$ (i.e., all possible outcomes are accounted for), then for any random variable X (with outcomes in the domain specified by the B_j):*

$$Pr[X] = \sum_{j \in [1,n]} Pr[X \mid B_j] Pr[B_j]$$

A common application of this law is when there is some event E that either happens or does not happen. In that case, for any random variable X:

$$Pr[X] = Pr[X \mid E] Pr[E] + Pr[X \mid \not E] Pr[\not E]$$

This yields the following useful bound:

$$Pr[X \mid E] Pr[E] \leq Pr[X] \leq Pr[X \mid E] Pr[E] + Pr[\not E]$$

Fact 9 (Law of Total Expectation) *For any random variables X, Y:*

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X \mid Y]]$$

This is related to the following fact that follows from the law of total probability: Given any collection of disjoint possible outcomes $B_1, X_2, \ldots B_n$ where $\sum_{j=1}^{n} Pr[B_j] = 1$ (i.e., all possible outcomes are accounted for), then for any random variable X (with outcomes in the domain specified by the B_j):

$$\mathrm{E}[X] = \sum_{j \in [1,n]} \mathrm{E}[X \mid B_j] \, Pr[B_j]$$

5 Standard tail bounds

In this section, I will list some of the important tail bounds that we have used repeatedly throughout the semester so far. These are immensely useful tools when you want to understand the performance of a randomized algorithm.

Theorem 10 (Markov's Inequality) *For any non-negative random variable X, for any value a:*

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

Of note: $Pr[X \geq 2\mathrm{E}[X]] \leq 1/2$.

Theorem 11 (Chebychev's Inequality) *For any non-negative random variable X:*

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

Theorem 12 (Chernoff Bound (upper tail)) *For any collection of **independent** random variables $X_1, X_2, \ldots, X_n$ where each $X_j \in [0,1]$: Let $X = \sum_{j=1}^{n} X_j$ and $\mu = \mathrm{E}[X]$. Fix a positive constant δ.*

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

As an application of this, we observe that:

$$Pr[X \geq e\mu] \leq e^{-\mu}$$

Alternatively, we can simplify the bound:

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

For $\delta \leq 1$, we get:

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

Theorem 13 (Chernoff Bound (lower tail)) *For any collection of **independent** random variables $X_1, X_2, \ldots, X_n$ where each $X_j \in [0,1]$: Let $X = \sum_{j=1}^{n} X_j$ and $\mu = \mathrm{E}[X]$. Fix a positive constant $\delta \leq 1$.*

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

Alternatively, we can simplify the bound:

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

Theorem 14 (Hoeffding Bound) *For any collection of **independent** random variables $X_1, X_2, \ldots X_n$ where each $X_j \in [0,1]$: Let $X = \sum_{j=1}^{n} X_j$ and $\mu = \mathrm{E}[X]$. For any value t:*

$$Pr[|X - \mu| \geq t] \leq 2e^{-2t^2/n}$$

6 A few useful mathematical inequalities

Below are a few useful mathematical facts that have appeared in some of the algorithm analysis we have done. These are standard facts and not particularly specific to randomized algorithms. But they are useful!

The following fact follows from the Taylor expansion of the function $e^x = 1 + x + x^2/2 + \ldots$:

Fact 15 *If* $0 < x \leq 1$*:* $1 - x \leq e^{-x} \leq 1 - x/2$.

From this we immediately derive the following fact.

Fact 16 *If* $0 < p \leq 1$*:* $e^{-2} \leq (1-p)^{1/p} \leq e^{-1}$.

Notice that we often use this fact when $p = 1/n$, concluding that $(1 - 1/n)^n \leq 1/e$.

It is often useful to be able to approximate $\binom{a}{b}$, i.e., the number of ways to choose a items from a set of b items:

Fact 17 $\left(\frac{a}{b}\right)^b \leq \binom{a}{b} \leq \left(\frac{ea}{b}\right)^b$.

There are also a variety of standard summations that are useful.

Fact 18 *If* $0 < \alpha < 1$*:* $\sum_{i=0}^{\infty} \alpha^i = \frac{1}{1-\alpha}$

Fact 19 $\ln(n-1) \leq \sum_{i=1}^{n} \frac{1}{i} \leq \ln(n) + 1$

Fact 20 $\sum_{i=0}^{j} 2^i = 2^{j+1} - 1$