Algorithms, 2023 Fall, Homework 11
(Due Dec 13)

December 4, 2023

Problem 1 : () KT book pp 416, problem 5

Problem 2 : () KT book pp 416, problem 6

Problem 3 : KT book pp 420, problem 12

Problem 4 : (Filling out a 0/1 matrix) You are asked to construct an $n \times n$ $0/1$ matrix (i.e., each cell is either 0 or 1) that satisfies the following set of properties: The ith row contains r_i "1"s for all $1 \leq i \leq n$. The jth column contains c_j "1"s for all $1 \leq j \leq n$. The integers r_is and c_js are given and such that $\sum_{i=1}^{n} r_i = \sum_{j=1}^{n} c_j$. A *permutation matrix* matrix is such $0/1$ matrix with $r_i = c_i = 1$ for all $1 \leq i \leq n$.

1. Suppose M is an $n \times n$ $0/1$ matrix such that $r_i = c_i = \ell$ for all $1 \leq i \leq n$ and some positive integer $\ell \leq n$. Prove any such $0/1$ matrix can be write as a sum of ℓ permutation matrices.

2. Suppose all r_is and c_js are even numbers. You are also given an $n \times n$ matrix M that satisfies the properties. Show there exists an $n \times n$ matrix M' that satisfies the properties with respect to $r_i/2$ and $c_j/2$ for all i, j. Moreover, M' must satisfy that $M_{ij} = 1$ if $M'_{ij} = 1$ (i.e., M' is obtained from M by changing some 1s to 0s). (hint: you may want to use the fact I mentioned in the class: there is an integral max flow if all edge capacities are integers.)

Problem 5 : Suppose we want to solve a minimization problem. Assume we have an randomized approximation algorithm which can find in polynomial time a (random) solution with an expected value $E \leq \alpha OPT$ where OPT is the optimal minimum value. Show that there is a polynomial time algorithm that can find a good solution with high probability. More precisely, for any small $\epsilon, \delta > 0$, there is an randomized algorithm that produces a solution S such that

$$\Pr[S < (1+\epsilon)\alpha OPT] \geq 1 - \delta$$

(The running time of the algorithm may depends on ϵ and δ.)

Problem 6 : Derandomize the randomized 2-approximation algorithm for max-cut. Write down all the details. Interpret the algorithm as a greedy algorithm.

Problem 7 : Y is a random variable with mean μ. We do not know the distribution of Y. However, we can get samples drawn from the distribution of Y. We also know that $\Pr[|Y-\mu| > k] \leq 1/4$. For any $\delta > 0$, show that it is possible to obtain an estimate ν of μ, such that $\Pr[|\nu - \mu| > k] \leq \delta$ by using only $O(\log \frac{1}{\delta})$ i.i.d. samples. (hint: there is a very simple solution). This is a very useful idea to boost the success probability.

Problem 8 : Suppose we sample n points uniformly from $[0, 1]$. It seems (to some) that the lengths of all $n+1$ segments are distributed equally. However, some students would say, wait, they are not the same to me, as the pieces in the boundary may not be distributed in the same way as the one in the middle. Some came up with the following argument: a circle is completely symmetric. Let us say we sample n points uniformly from this circle. Hence, each piece is completely symmetric and hence distributed equally. Now, it is safe to conclude the expected length of each piece is $1/n$. Now, we choose a point that break the circle to a line segment (the sampling does not need to know where this breaking point was). But now the first and last segments add up to one segment in the circle, so each of them has expected length $1/2n$.

Which argument is correct? Which part is wrong with the wrong argument (explain it using math)? Or if you think none of the above is correct, you should provide your own correct statement.

(hint: start with $n = 2, 3$ and compute the distribution of each piece rigorously. How you see how seemingly correct argument can be awfully wrong, if one is not completely rigorous)

Bonus problems: Bonus problems are problems that may be (may be not) more challenging than other problems. You do not lose points if you do not answer bonus problems. But you will get extra points if you get correct answers for such problems. **To make your life easier, we give you two weeks for the bonus problems**).

Problem 1 : (You may need Chernoff bound for the following problem.) We know what is an FPTAS. An FPRAS (fully polynomial time randomized approximation scheme) is the randomized version of FPTAS. A *fully polynomial randomized approximation scheme (FPRAS)* for an estimation problem f (e.g., estimating the number of perfect matchings in a graph, or a value of an integral formula) is a randomized algorithm A that takes an input instance x a real number $\epsilon > 0$, returns $A(x)$ such that $\Pr[(1-\epsilon)f(x) \leq A(x) \leq (1+\epsilon)f(x)] \geq \frac{3}{4}$ and its running time is polynomially in both the size of the input n and $1/\epsilon$.

Assume that for any convex body K in $\mathbb{R}^n$, we can

1. sample uniformly from K in time polynomial of n;
2. compute its volume $\mathrm{vol}(K)$ precisely in time polynomial of n.

Consider the following problem: Suppose that we have m convex bodies $K_1, K_2, \ldots, K_m$, in n dimensions, provided by m membership oracles $\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_m$, where for any n-dimensional point x, and any $1 \leq i \leq m$, $\mathcal{O}_i(x)$ can decide whether $x \in K_i$ in constant time. With the above ideal assumptions and inputs: Give an FPRAS for estimating

$$\mathrm{vol}\left(\bigcup_{i=1}^{m} K_i\right).$$