

Problem Set 3

Out: 21 Oct.; Due: 4 Nov.

Notes: Solutions are due by **5pm on Friday November 4**. Please submit your solutions on Gradescope by that time; **remember to begin each problem (not problem part) on a new page**. Late solutions will not be accepted. Solutions should preferably be typeset in \LaTeX ; if this poses a problem then they may be written **neatly** by hand and scanned. Depending on grading resources, we reserve the right to grade only a subset of the problems and check off the rest (but since you don't know which subset, you're strongly advised to do all of them!).

Take time to write **clear** and **concise** answers. None of the problems require a long solution (often not much longer than the problem statement itself!); if you find yourself writing a lot, you are either on the wrong track, confused, or giving too much detail. You are actively encouraged to form small groups (two or three people) to work through the problems, but you **must always** write up your solutions on your own. If you use external sources, you should cite them; and again, you must understand and formulate your solutions yourself.

1. Another unbiased estimator for the permanent

Let A be an $n \times n$ 0-1 matrix. Consider the following algorithm for estimating $\text{per}(A)$, the permanent of A :

let B be a random matrix obtained by replacing each 1-entry of A with ± 1
 chosen independently and uniformly at random
 output the estimator $X_A = (\det(B))^2$

- (a) Show that X_A is an unbiased estimator of $\text{per}(A)$.
- (b) Show that the estimator performs very badly in the worst case by giving a specific family of matrices on which its variance is exponentially large. (Your family should have the property that X_A is zero with very high probability.)
- (c) For a 0-1 matrix A , let G_A denote the corresponding bipartite graph, and define the random variable $\gamma(A) = \mathbb{E}[3^{c(M, M')}]$, where the expectation is taken over all ordered pairs (with replacement) of perfect matchings (M, M') in G_A , and $c(M, M')$ is the number of cycles formed by the union of M and M' . (An edge shared by M, M' does not count as a cycle.) Show that the “critical ratio” of the above estimator is given by

$$\frac{\mathbb{E}(X_A^2)}{\mathbb{E}(X_A)^2} = \gamma(A).$$

- (d) We will show that, for a random matrix A , $\gamma(A)$ is polynomially bounded with high probability. Thus the above estimator, like Rasmussen's estimator in Lecture 12, also provides a FPRAS for the permanent of a random matrix. Once again, the key ingredient in the analysis is the concentration of the permanent of a random $n \times n$ matrix with m ones (Lemma 12.10). Use part (ii) of that lemma and Chebyshev's inequality to show that, assuming as usual that $\frac{m^2}{n^3} \rightarrow \infty$,

$$\Pr_{\mathcal{A}_{n,m}}[\text{per}(A)^2 \leq \frac{1}{2} \mathbb{E}_{\mathcal{A}_{n,m}}(\text{per}(A)^2)] = O\left(\frac{n^3}{m^2}\right).$$

[HINT: This is slightly stronger than what we showed at the bottom of page 12-5 with a single application of Chebyshev's inequality, but you can prove it using part (ii) of the lemma twice.]

- (e) To complete the analysis, we would like to compute the expected value of $\gamma(A)$ for a random matrix A . Unfortunately this is hard to do directly. However, we can compute expectations in a slightly different probability space. Let Ω denote the set of all triples of the form (A, M, M') , where $A \in \mathcal{A}_{n,m}$ is a matrix and (M, M') is a pair of perfect matchings in G_A , and assume that triples (A, M, M') are chosen uniformly at random from Ω . (Note that in this probability space, a matrix A appears with probability proportional to $\text{per}(A)^2$.) By counting permutations in similar fashion to the proof of Lemma 12.10, it can be shown that, in this probability space, the expected value of $3^{c(M, M')}$ is at most Cn^2 , for a specific constant C , again assuming that $\frac{m^2}{n^3} \rightarrow \infty$. Use this fact, together with part (d), to show that

$$\Pr_{\mathcal{A}_{n,m}}[\gamma(A) \geq n^2\omega(n)] \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where $\omega(n)$ is any function that goes to infinity with n . [HINT: Write the expectation of $3^{c(M, M')}$ under the uniform distribution over Ω as $\frac{1}{|\Omega|} \sum_A (\text{per}(A))^2 \gamma(A)$. Then assume that the probability above does not tend to 0 and obtain a contradiction.]

- (f) Finally, deduce from part (e) that the estimator X_A gives a FPRAS for the permanent of a random $n \times n$ matrix A with probability tending to 1 as $n \rightarrow \infty$.

2. Chernoff for Poisson

Let X be a Poisson random variable with parameter μ .

- (a) Derive the following Chernoff bound on the upper tail of X :

$$\Pr[X \geq \mu + \lambda] \leq \exp\{-(\mu + \lambda) \ln \frac{\mu + \lambda}{\mu} - \lambda\}.$$

Also, state and prove the analogous bound for the lower tail $\Pr[X \leq \mu - \lambda]$. [HINT: You should be able to explicitly compute $E[e^{Xt}]$ here.]

- (b) Deduce that X satisfies

$$\Pr[X \geq (1 + \beta)\mu] \leq \exp\{-\mu(-\beta + (1 + \beta) \ln(1 + \beta))\}$$

for any $\beta > 0$, which is exactly the Angluin bound for the binomial distribution given in Corollary 13.3 of Lecture 13. Also, derive the analogous bound for the lower tail.

3. Random geometric graphs

In a popular model for wireless networks, n points are placed independently and uniformly at random in a unit square, and each point is connected by an edge to the k other points that are closest to it. Call the resulting undirected graph G . The idea is that the points represent sensors, and each sensor is able to communicate with the k other sensors that are closest to it in space. In this problem we are going to prove that, in a weak sense, $k = \Theta(\log n)$ is a threshold for the graph G to be connected.

- (a) Show that if $k = c_1 \log n$ for a sufficiently large constant c_1 , then $\Pr[G \text{ is connected}] \rightarrow 1$ as $n \rightarrow \infty$. [HINT: Partition the unit square into small squares of area $\frac{\log n}{n}$; here and elsewhere you should ignore issues of rounding to integer values. Use a union bound to argue that whp every square contains at least one point. Use a Chernoff bound to show that the number of points within radius (say) $\sqrt{10 \frac{\log n}{n}}$ of any given point is at most $c \log n$ for some constant c whp. Deduce that whp every point is connected to all points within its own square and within all neighboring squares.]
- (b) Now assume conversely that $k = c_2 \log n$ for a sufficiently small constant c_2 (to be chosen later), and let r be such that $\pi r^2 = \frac{k+1}{n}$. Consider a set of three concentric discs D_1, D_3, D_5 of radii $r, 3r, 5r$ respectively such that D_5 is contained in the unit square. Call the set of discs “bad” if (i) D_1 contains at least $k+1$ points; (ii) $D_3 \setminus D_1$ contains no points; and (iii) the intersection of $D_5 \setminus D_3$ with each disc of radius $1.5r$ centered at a set of points spaced evenly at distance $0.01r$ around the boundary of D_3 contains at least $k+1$ points. Show that if our point set contains a bad set of discs, then the resulting graph G is disconnected.

- (c) For our final calculation in part (d) below, it will be convenient to assume that the points are distributed in the unit square according to a “Poisson Point Process” (PPP) of intensity n . In this model, the number of points in any subregion A of the square has a Poisson distribution with parameter $n \times \text{area}(A)$, and the numbers of points in disjoint subregions are independent. (This independence makes things much simpler.) An *occupancy event* is an event that depends only on the number of points in some region of the unit square. Suppose an occupancy event \mathcal{E} holds in the PPP model with probability at most $\delta(n)$, where $\delta(n) \ll 1/\sqrt{n}$. Show that \mathcal{E} holds in the original n -point model with probability tending to zero as $n \rightarrow \infty$. [HINT: Recall from Lecture 15 the relationship between the standard balls-and-bins model and its independent Poisson analog. You may use the fact that, for a Poisson(λ) r.v. X where λ is a non-negative integer, $\Pr[X = \lambda] \geq \frac{1}{e\sqrt{\lambda}}$.]
- (d) Show that, by choosing $k = c_2 \log n$ for a sufficiently small constant c_2 , we can ensure that the probability in the PPP model that any particular set of three discs as in part (b) is bad is at least $\Omega(n^{-1+\epsilon})$ for some $\epsilon > 0$. Hence deduce that, in the PPP model, the probability that a bad set of discs exists is at least $1 - \exp(-cn^\epsilon / \log n)$ for some constant $c > 0$. [HINT: Use the Chernoff bound for Poisson r.v.’s in Q2(b) above.]
- (e) Finally, put parts (b), (c), (d) together to deduce that, in the original n -point model, if we take $k = c_2 \log n$ with a sufficiently small constant c_2 , we have $\Pr[G \text{ is connected}] \rightarrow 0$ as $n \rightarrow \infty$.

[NOTE: With more effort, a stronger threshold property can be proved here, namely there exists a constant c_0 such that if $k = c \log n$ then $\Pr[G \text{ is connected}] \rightarrow 1$ if $c > c_0$ and $\Pr[G \text{ is connected}] \rightarrow 0$ if $c < c_0$.]

4. Codes in space

This question concerns the randomized construction of codes with certain spatial properties, which are useful in certain geometric hardness of approximation arguments.

- (a) An (m, ℓ, ϵ) -code is a family of m strings (codewords) of length ℓ over a fixed alphabet $\{1, 2, \dots, a\}$ such that any pair of distinct strings agree in at most $\epsilon\ell$ positions. Let \mathcal{S} be a set of m strings of length ℓ chosen independently and u.a.r. Show that, for any alphabet size $a > \frac{1}{\epsilon}$, if $\ell \geq C \ln m$ for a suitable constant C (depending on a and ϵ), then the strings form an (m, ℓ, ϵ) -code with probability $\rightarrow 1$ as $m \rightarrow \infty$. [HINT: Use Chernoff to bound the probability of $\epsilon\ell$ collisions for any particular pair.]
- (b) Suppose we want to reduce the word length ℓ to $\delta \ln m$ for an arbitrary constant $\delta > 0$. Show that we can do this by taking the alphabet size a to be a sufficiently large constant (which depends on ϵ and δ). [HINT: You will need the following sharper form of Chernoff/Hoeffding bound from Lecture 13: $\Pr[X \geq (1 + \beta)\mu] \leq \exp\{-\mu((1 + \beta) \ln(1 + \beta) - \beta)\}$.]
- (c) A (3-dimensional) (m, ℓ, ϵ) -grid code is defined similarly to the above except that no pair of strings may “score” more than $\epsilon\ell$ in any embedding of the two strings on a 3-dimensional rectangular grid. An *embedding* of one string is an assignment of its symbols to the consecutive vertices of a simple path (i.e., no loops) in the grid; an embedding of a *pair* of strings is an embedding of each string such that the two corresponding paths are vertex-disjoint (i.e., the strings do not touch each other). The *score* of

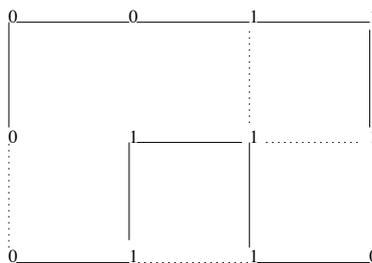


Figure 1: An embedding of the pair of strings 000111 and 011110 in the two-dimensional grid. The score of this embedding is 4; the scoring symbol pairs are indicated by dotted lines.

a pair of embedded strings is the number of pairs of identical symbols that lie at adjacent grid points in the embedding (excluding those pairs that are consecutive in one or other of the strings, and hence are always adjacent in any embedding). See Fig. 1 for a 2-dimensional example.

Show that the same random construction, with word length $\ell = \delta \ln m$, gives an (m, ℓ, ϵ) -grid code with probability $\rightarrow 1$ as $m \rightarrow \infty$, provided the alphabet size a is a large enough constant (depending on ϵ and δ). [HINT: You should start by showing that the number of possible sets of adjacencies that can arise between any pair of adjacent strings of length ℓ is at most $\exp(\alpha\ell)$ for some constant α : there is no need to optimize α , a very crude bound will suffice. Then use the form of Chernoff in part (b) to bound the probability of a large deviation from the mean score in any one embedding. Unfortunately, there is a little catch: the events you are counting are *not* independent—why?—but you can get around this by breaking them up into three subsets of independent events, one for each dimension.]

5. More on the power of two choices

In Lecture 15 we saw that, if n balls are thrown into n bins under the random two-choice scheme, the maximum load in any bin is at most $\frac{\ln \ln n}{\ln 2} + O(1)$ with high probability. Prove a matching lower bound, i.e., show that the maximum load is at least $\frac{\ln \ln n}{\ln 2} + \Omega(1)$ with high probability. [HINT: In similar fashion to the sequence β_i in the proof of the upper bound in Lecture 15, define a sequence α_i with the following property: after $(1 - 1/2^i)n$ balls have been thrown, the number of bins with at least i balls is at least α_i with high probability. (I.e., consider throwing the first $n/2$ balls, then the next $n/4$, the next $n/8$, etc.) The sequence α_i should be decreasing, with $\frac{\alpha_{i+1}}{n} \ll \left(\frac{\alpha_i}{n}\right)^2$; for example, you might find it convenient to set $\alpha_{i+1} = \frac{1}{2^{i+3}} \frac{\alpha_i^2}{n}$. Use iterated Chernoff bounds as in the upper bound proof.]