Liam Hardiman *February 15, 2021*

Math 274 - Homework 2

Problem 1. Let $v_1 = (x_1, y_1), \dots, v_n = (x_n, y_n)$ be n vectors in $\mathbb{Z}^2$, where each x_i and each y_i is a positive integer that does not exceed $\frac{2^{n/2}}{10\sqrt{n}}$. Show that there exist two disjoint nonempty subsets $I, J \subseteq [n]$ such that $\sum_{i\in I} v_i = \sum_{j \in J} v_j$.

Proof. Consider the random sum $V = (X, Y) = \sum_{i=1}^n \epsilon_i v_i$ where each ϵ_i is an iid Bernoulli random variable with success probability $1/2$. We'll show that a sizeable proportion of the possible V live in an axis-aligned box centered about the mean of V. If the size of this box is smaller than the number of assignments of the ϵ_i's that make V land in this box, then there must be two assignments of the ϵ_i's that give the same realization of V.

Since X is a sum of independent scaled Bernoulli random variables, its variance is easily computed:

$$\mathrm{Var}[X] = \sum_{i=1}^n x_i^2 \cdot \mathrm{Var}[\epsilon_i] \leq \frac{2^n}{400}.$$

Note that the y-coordinate has the same variance. Now by Chebyshev we have

$$\Pr\left[\left|X - \frac{2^n}{400}\right| \leq t\right] \geq 1 - \frac{\mathrm{Var}[X]}{t^2},$$

and the same inequality holds for Y. If we choose t so that both of these probabilities are at least $3/4$, then V will leave in an axis-aligned square centered at $\mathbb{E}[V]$ with width $2t$ with probability at least $1/2$. Choosing $t = 2^{n/2}/10$ does precisely this. The resulting box contains $\sim 2^n/25$ integer points as well as at least 2^{n-1} possible realizations of V. Since there are more realizations than integer points in this box, two realizations must land on the same integer point. In particular, there are some *nonempty* and *distinct* $I', J' \subseteq [n]$ such that $\sum_{i\in I'} v_i = \sum_{j\in J'} v_j$. Setting $I = I' \setminus J'$ and $J = J' \setminus I'$ gives us the desired *disjoint* sets. □

Problem 2. Let $G = (V, E)$ be a graph on n vertices and let $0 < \alpha < 1$. Consider the following game on the edge set of G. There are two players, Alice and Bob, who alternate turns (starting with Alice) occupying perviously unoccupied edges in E until there are no more edges to claim. Let $G_A = (V, E_A)$ be the spanning subgraph of G where E_A is the set of edges occupied by Alice at the game's conclusion. Alice wins if and only if for every vertex $v \in V$, its degree in G_A is at least an α-fraction of its original degree in G.

Show that for large enough n, if G has minimum degree at least $\log^2 n$, then Alice has a winning strategy for $\alpha \geq 1/3 - \epsilon$.

Proof. Consider the following randomized strategy for Alice. She opens by taking any random edge in E. On each of her subsequent turns, she examines the edge $\{x, y\}$ that Bob just claimed. She flips a fair coin – claiming a random unclaimed edge incident to x if it comes up heads and incident to y if tails. If neither x nor y has any unclaimed incident edges, Alice simply picks up a random unclaimed edge in E.

If Alice wins with this strategy with positive probability, then any realization of winning coin flips corresponds to a *deterministically* winning strategy. Now for any vertex v in V, let $d_A(v)$ denote its degree in G_A and $d(v)$ its degree in G. The expected value of $d_A(v)$ is at least $\frac{1}{3}d(v)$ since, on average, Alice claims an edge incident to v at least half as often as Bob does. By Chernoff, for any $\epsilon < 1$ and any vertex v we have

$$\Pr[d_1(v) < (1-\epsilon)d(v)/3] \leq e^{-\epsilon^2 d(v)/6} \leq n^{-\frac{\epsilon^2}{6}\log n}.$$

A union bound then gives

$$\Pr[d_1(v) < (1-\epsilon)d(v)/3 \text{ for some } v] \leq n^{1-\frac{\epsilon^2}{6}\log n}.$$

Alice loses if she claims less than an α-fraction of the neighbors of any vertex v and solving the equation $1 - epsilon = 3\alpha$ for α lets us conclude that she wins with positive probability when $\alpha = 1/3 - \epsilon$. ☐

Problem 3. Prove that for every constant d, there exists a constant C so that for any given graph G with $|V(G)| = n$ (where n is arbitrarily large) and $\Delta(G) \leq d$, there exists an injective function $f : V \to [Cn]$ such that the quantities $\{|f(u) - f(v)| : uv \in E(G)\}$ are all distinct.

Proof. Choose f uniformly at random among all injections $V \to [Cn]$ and for any pair of edges uv and xy in E, define $A_{uv,xy}$ to be the event that $|f(u) - f(v)| = |f(x) - f(y)|$. Furthermore, let $N(uv, xy)$ denote the set of (unordered) pairs of edges, one of which is equal to either uv or xy. We claim that for any pair of edges uv, xy and any $J \subseteq \binom{E}{2} \setminus N(uv, xy)$, we have

$$\Pr\left[A_{uv,xy} \,\middle|\, \bigcap_{(u'v',x'y') \in J} A^c_{u'v',x'y'}\right] \leq \Pr[A_{uv,xy}].$$

Indeed, set E_J to be the event being conditioned on in the above expression for ease of notation. We have

$$\Pr[A_{uv,xy} \mid E_J] = \sum_{j=1}^{n-1} \Pr\Big[|f(u) - f(v)| = |f(x) - f(y)| = j \mid E_J\Big].$$

Informally, since f is an injection, each $j = 1, \ldots, n-1$ can be realized as a distance $|f(u) - f(v)|$ only so many times (depending on j). Each $A^c_{u'v',x'y'}$ hidden behind E_J represents a pair of edges, disjoint from uv, xy attaining two different distances under f. Since $u'v'$ and $x'y'$ attain different distances under f, it is more likely that at least one of them is j, and consequently, it is less likely that uv and xy are both separated by a distance of j under f. Thus, the above quantity is at most $\sum_{j=1}^{n-1} \Pr[|f(u) - f(v)| = |f(x) - f(y)| = j] = \Pr[A_{uv,xy}]$.

We have then established that the $A_{uv,xy}$ and $N(uv, xy)$ form a negative dependency graph. Now $|N(uv, xy)|$ is the number of edge pairs, one of which is equal to uv or xy. There are at most $2(m-1) \leq$

$n\Delta - 2$ such pairs. Let's bound $\Pr[A_{uv,xy}]$ from above. If uv and xy have a vertex in common, they are less likely to satisfy $|f(u) - f(v)| = |f(x) - f(y)|$ than if they had no vertex in common. We then suppose that uv and xy share no vertex. We have

$$\begin{aligned}
\Pr[A_{uv,xy}] &= \sum_{j=1}^{n-1} \Pr\left[|f(u) - f(v)| = |f(x) - f(y)| = j\right] \\
&\leq \sum_{j=1}^{n-1} \frac{4 \cdot \frac{(n-j)^2}{2} \cdot \binom{cn-4}{n-4}(n-4)!}{\binom{cn}{n} n!} \\
&\sim \frac{2}{n^4} \cdot \frac{\binom{cn-4}{n-4}}{\binom{cn}{n}} \cdot \sum_{j=1}^{n-1} (n-j)^2 \\
&\sim \frac{2}{3nc^4}.
\end{aligned}$$

Call this (asymptotic) upper bound p. Since $ep(n\Delta - 1) \leq \frac{2e\Delta}{3c^4}$, we may simply choose c large enough so that this quantity is less than 1 and apply the Lovász local lemma to conclude that the distances $|f(u) - f(v)|$ are all distinct with positive probability (and large enough n). $\square$

Problem 4. Prove that for n sufficiently large, any set A of n distinct integers contains two disjoint subsets B_1 and B_2 satisfying $|B_1| = |B_2| > 0.33n$, where each of the sets B_i is sum-free.

Proof. Let q be a prime much larger than the largest element of A, say congruent to 1 modulo 6, so $q = 6k+1$ for some k. Notice that the interval $S_1 = \{2k+1, 2k+2, \ldots, 4k\}$ is sum-free modulo q, as is the union of intervals $S_2 = \{k+1, \ldots, 2k-1\} \cup \{4k+1, \ldots, 5k\}$. Consider the transformation $T : A \to \mathbb{Z}_q$ where $a \mapsto ax + y \pmod{q}$ where x and y are nonzero integers modulo q chosen independently at random. Notice that for any $u, v \in \mathbb{Z}_q$, and any distinct a_1, a_2 in A, the events $\{T(a_1) = u\}$ and $\{T(a_2) = v\}$ are independent. Indeed, these events each occur with probability $1/q$ and

$$\Pr[a_1 x + y \equiv u,\ a_2 x + y \equiv v] = \Pr\left[\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \equiv \begin{pmatrix} u \\ v \end{pmatrix}\right] = \frac{1}{q^2},$$

since the above matrix has rank 2 (so there is a *unique* pair x, y solving the corresponding system).

Now $B \subseteq A$ is sum-free if $T[B] \subseteq (S_1 + y)$. Indeed, if $b_1 + b_2 = b_3$ for b_1, b_2, b_3 in B, then $b_i x = s_i$ for some s_i in S_1 for $i = 1, 2, 3$ and $s_1 + s_2 = s_3$, contradicting the fact that S_1 is sum-free. The same holds for S_2. Now if we can show that the expected number of a_i such that $T(a_i) \in (S_1 + y)$ is at least $n/3$ (and that the same holds for S_2), the independence of the events $\{T(a_i) \in (S_1 + y)\}$ and $\{T(a_j) \in (S_2 + y)\}$ for distinct a_i and a_j will guarantee the existence of disjoint sum-free subsets B_1 and B_2, both of size around 1/3.

For any interval I of length ϵn in $\mathbb{Z}_q$, the expected size of $T[A] \cap I$ is $\epsilon n^2/q$. If we write $|T[A] \cap I| = \sum_{a \in A} X_{a,I}$, where $X_{a,I}$ indicates whether $T(a) \in I$, then we have

$$\mathrm{Var}\Big[|T[A] \cap I|\Big] = \sum_{a \in A} \mathrm{Var}[X_{a,I}] = n \cdot \frac{\epsilon n^2}{q} \cdot \left(1 - \frac{\epsilon n^2}{q}\right),$$

since the $X_{a,I}$ are independent by our construction of T from *two* random sources x and y. By Chebyshev's inequality, we have

$$\Pr\Big[\big||T[A] \cap I| - \epsilon^2 n/q| \geq \epsilon^2 n/q\big] = o(1),$$

so for n sufficiently large, a $1 - o(1)$ fraction of the intervals of $\mathbb{Z}_q$ contain the expected number of elements of $T[A]$. Consequently, $|T[A] \cap (S_1 + y)|$ and $|T[A] \cap (S_2 + y)|$ are both simultaneously near their expected sizes, say equal to $0.33n$. □

Problem 5. Prove that there exists an absolute constant $c > 0$ such that for any tournament $T = (V, E)$ on n vertices, there are two disjoint subsets $A, B \subseteq V$ such that $e(A, B) - e(B, A) \geq cn^{3/2}$.

Proof. I think the idea here is to choose A and B in some random way and then show that $\mathbb{E}[(e(A, B) - e(B, A))^2] \geq cn^3$. There would then be some realization of A and B for which $(e(A, B) - e(B, A))^2$ is at least cn^3 and we'd be done. I wasn't able to come up with a way to get a variance with order higher than n^2, unfortunately.

Intuitively, I wanted two random sets of roughly equal size, so for each element in V, we include it in A with probability p, in B also with probability p, and neither with probability $1 - 2p$. Since $e(A, B)$ and $e(B, A)$ have the same distribution, we have

$$\mathbb{E}[(e(A, B) - e(B, A))^2] = 2\mathbb{E}[e(A, B)^2] - 2\mathbb{E}[e(A, B)e(B, A)].$$

For each edge e we let X_e be the variable that indicates whether e crosses from A to B. We similarly let $\tilde{X}_e$ indicate whether e crosses from B to A. We have

$$\mathbb{E}[e(A, B)^2] = \mathbb{E}\left[\left(\sum_{e \in E} X_e\right)^2\right] = \binom{n}{2} p^2 + \sum_{e \neq f} \mathbb{E}[X_e X_f].$$

If the edges e and f share no common vertex, then $\mathbb{E}[X_e X_f] = p^4$. If both are in- or out-edges to a common vertex, then $\mathbb{E}[X_e X_f] = p^3$. If one is an in-edge and one is an out-edge, then $\mathbb{E}[X_e X_f] = 0$. Now we compute

$$\mathbb{E}[e(A, B)e(B, A)] = \mathbb{E}\left[\left(\sum_{e \in E} X_e\right)\left(\sum_{f \in E} \tilde{X}_f\right)\right] = \sum_{e, f \in E} \mathbb{E}[X_e \tilde{X}_f].$$

If $e = f$ then we clearly have $\mathbb{E}[X_e \tilde{X}_f] = 0$. If they don't share a vertex, then we get p^4. When they're both in- or out-edges of some vertex then we get zero, and if one is an in-edge and the other an out-edge, we get p^3. Let S_+ be the set of all edge pairs that are out-edges to some common vertex, S_- to be the pairs that form in-edges, and $S_\pm$ the pairs where one is an in-edge and the other an out-edge. Since the contributions due to edge pairs without common vertices cancel with each other in our expressions for $\mathbb{E}[e(A,B)^2]$ and $\mathbb{E}[e(A,B)e(B,A)]$, we have

$$\mathbb{E}[(e(A,B)-e(B,A))^2] = p^2\binom{n}{2} + 2p^3\Big(|S_-|+|S_+|-|S_\pm|\Big).$$

We can exactly calculate the expression in parentheses above by summing over each vertex ($d^+(v)$ and $d^-(v)$ are the out- and in- degrees of v, respectively)

$$\begin{aligned}
|S_| + |S_+| - |S_\pm| &= \sum_{v\in V}\left[\binom{d^-(v)}{2} + \binom{d^+(v)}{2} - d^-(v)d^+(v)\right] \\
&= \frac{1}{2}\sum_{v\in V}(d^+(v)-d^-(v))^2 - \frac{1}{2}\sum_{v\in V}(d^-(v)+d^+(v)) \\
&= \frac{1}{2}\sum_{v\in V}(d^+(v)-d^-(v))^2 - \binom{n}{2}.
\end{aligned}$$

Our second moment is then

$$\mathbb{E}[(e(A,B)-e(B,A))^2] = \binom{n}{2}(p^2-2p^3) + p^3\sum_{v\in V}(d^+(v)-d^-(v))^2.$$

By considering the case of a regular tournament (wherein $d^+(v)=d^-(v)$ for all $v\in V$), we see that the above quantity can't be made larger than $\Theta(n^2)$ in general. Interestingly, if we set $p=1/2$ and consider a regular tournament, we obtain a variance of zero, which concludes a circuitous proof for the fact that $e(A,A^c)-e(A^c,A)=0$ for any subset A in this case.

I must have the wrong experiment for choosing A and B. This regular tournament example tells me that we can't do something as simple as take $B=A^c$. I think setting $p=1/3$ above chooses disjoint subsets A and B of V uniformly at random and this still doesn't work. □