2020 FALL MAS583 HW2

Fanchen Bu [20194185]

September 17, 2020

1 Problem 1

1.1 Question

Suppose $n \geq 2$ and let $H = (V, E)$ be an n-uniform hypergraph with $|E| = 4^{n-1}$ edges. Show that there is a coloring of V by four colors so that no edge is monochromatic.

1.2 Answer

Let X be the random variable denoting the number of monochromatic edges in H. Clearly,

$$X = \sum_{e \in E} X_e,$$

where X_e is the indicator random variable for edge e being monochromatic, i.e.,

$$X_e = \begin{cases} 1, & e \text{ is monochromatic} \\ 0, & \text{otherwise} \end{cases}$$

Then if we operate a coloring of V where we color each edge in four colors with equal probability (the probability of every color is $1/4$), randomly and uniformly, we can easily get $\mathbb{E}[X_e] = 1/4^{n-1}$, which implies $\mathbb{E}[X] = 1$.
Clearly, we can color H such that $X > 1$, e.g., we color all the edges in H in a single color. Thus, there also exist a coloring where $X < 1$, which means, as X is an integer, $X = 0$, completing the proof.

2 Problem 2

2.1 Question

Prove that there is a positive constant c so that every set A of n nonzero reals contains a subset $B \subset A$ of size $|B| \geq cn$ so that there are no $b_1, b_2, b_3, b_4 \in B$ satisfying

$$b_1 + 2b_2 = 2b_3 + 2b_4.$$

2.2 Answer

When $n < 2$, the proposition is trivial with any $c \leq 1$, so we only need to care about the cases when $n \geq 2$ and prove the proposition holds with some $c \leq 1$.
Let $A = \{a_1, a_2, ..., a_n\}$ be a set of $n \geq 2$ nonzero reals, and $\epsilon = \min_{i\in[n]} |a_i|$. We choose x randomly and uniformly in the interval $[1/\epsilon, 10n/\epsilon]$ and compute the numbers $d_i = xa_i \bmod 1$, $0 \leq d_i < 1$ for all $1 \leq i \leq n$. Let $D = \{d_1, d_2, ..., d_n\}$ and X be the random variable denoting the numbers of elements in D which are in the interval $[3/17, 4/17)$, i.e., $X = |D \cap [3/17, 4/17)|$. Clearly, $\mathbb{E}[X] = \sum_i \mathbb{E}[X_i]$, where X_i is the indicator random variable for d_i being in the interval $[3/17, 4/17)$. We can compute

$$\begin{aligned}\mathbb{E}[X_i] &\geq \frac{a_i(10n-1)/\epsilon - 2}{17a_i(10n-1)/\epsilon} \\ &\geq \frac{1}{17}(1 - \frac{2}{10n-1}).\end{aligned}$$

Thus,

$$\mathbb{E}[X] \geq \frac{1}{17}(n - \frac{2n}{10n-1}) \geq \frac{n}{34},$$

which implies that there is an x and a subsequence B of A with $|B| \geq n/34$ such that $xb \bmod 1 \in [3/17, 4/17)$ for each $b \in B$. We say a set B is 1222-sum-free if there are no $b_1, b_2, b_3, b_4 \in B$ satisfying $b_1 + 2b_2 = 2b_3 + 2b_4$. Then we claim the above B is 1222-sum-free since $[3/17, 4/17)$ is 1222-sum-free with respect to addition module 1 (easy to check, $b_1 + 2b_2 \in [9/17, 12/17)$ and $2b_3 + 2b_4 \in [12/17, 16/17)$), completing the proof.

3 Problem 3

3.1 Question

Prove that every set of n nonzero *real* numbers contains a subset A of *strictly* more than $n/3$ numbers such that there are no $a_1, a_2, a_3 \in A$ satisfying $a_1 + a_2 = a_3$.

3.2 Answer

Similarly, we can only care about cases when $n \geq 2$. And we will prove the proposition based on the theorem below[1], where we call a set A *sum-free* is there are no $a_1, a_2, a_3 \in A$ satisfying $a_1 + a_2 = a_3$.

Theorem 3.1 *Every set $B = \{b_1, ..., b_n\}$ of n nonzero integers contains a sum-free subset A of size $|A| > \frac{1}{3}n$.*

Let $B = \{b_1, b_2, ..., b_n\}$ be a set of $n \geq 2$ nonzero reals. If we can prove that there exist a set $C = \{c_1, c_2, ..., c_n\}$ of n nonzero integers and for any $i, j, k \in [n]$, $b_i + b_j = b_k$ if and only if $c_i + c_j = c_k$, then we can complete the proof.[2]

[1] Theorem 1.4.1 on the textbook.

[2] The original paper proved a similar proposition for sequences, which is, in my understanding, a stronger one, but we can construct a similar proof for sets.

Lemma 3.2 *Given any arbitrary set $B = \{b_1, b_2, ..., b_n\}$ of nonzero reals, there is a set $C = \{c_1, c_2, ..., c_n\}$ of nonzero integers such that, for any $\epsilon_1, ..., \epsilon_n \in \{2, 1, 0, -1, -2\}$, the sign of $\sum_{i=1}^n \epsilon_i b_i$ (the sign of a positive real is 1, the sign of a negative real is -1, and the sign of 0 is 0) is equal to that of $\sum_{i=1}^n \epsilon_i c_i$.*

Proof. For each of the 5^n possible vectors $\epsilon = (\epsilon_1, ..., \epsilon_n)$, let $E(\epsilon)$ be an equation or an inequality with the n variables $x_1, ..., x_n$ defined as follows: if $\sum_{i=1}^n \epsilon_i b_i = 0$ then $E(\epsilon)$ is the equation $\sum_{i=1}^n \epsilon_i x_i = 0$; if $\sum_{i=1}^n \epsilon_i b_i > 0$ then let q be a positive rational so that $\sum_{i=1}^n \epsilon_i b_i \geq q$ and let $E(\epsilon)$ be the inequality $\sum_{i=1}^n \epsilon_i x_i \geq q$; similarly if $\sum_{i=1}^n \epsilon_i b_i < 0$ then let q be a negative rational so that $\sum_{i=1}^n \epsilon_i b_i \leq q$ and let $E(\epsilon)$ be the inequality $\sum_{i=1}^n \epsilon_i x_i \leq q$. Consider the linear program in the n variables $x_1, ..., x_n$ consisting of the 3^n constraints $E(\epsilon)$. This program has a feasible real solution $(b_1, ..., b_n)$, where each b_i is distinct and nonzero. Since all the constraints have rational coefficients and the set of rational numbers is dense in $\mathbb{R}$, it also has a rational solution $(d_1, ..., d_n)$, where each d_i is distinct and nonzero. By multiplying all these numbers d_i by a suitable integer we obtain a desired set of integers $C = \{c_1, ..., c_n\}$. ■

With Lemma 3.2, for our B of n nonzero reals, there is a set C of n nonzero integers satisfying

$$\text{sign}(\sum_{i=1}^n \epsilon_i b_i) = \text{sign}(\sum_{i=1}^n \epsilon_i c_i),$$

for all $\epsilon_i \in \{2, 1, 0, -1, -2\}$. We can easily check what we want (for any $i, j, k \in [n]$, $b_i + b_j = b_k$ if and only if $c_i + c_j = c_k$) is satisfied as $b_i + b_j = b_k$ is equivalent to $b_i + b_j - b_k = 0$ or $\text{sign}(b_i + b_j - b_k) = 0$, completing the proof.

3.3 References

The main references for problem 2 and 3 are as follows:

- Alon, Noga, and Daniel J. Kleitman. Sum-free subsets. A tribute to Paul Erdos (1990): 13-26.

- Caro, Yair. Generalized sum-free subsets. International Journal of Mathematics and Mathematical Sciences 13 (1990).

- Cao, Zhengjun, and Lihua Liu. A Note on the Alon-Kleitman Argument for Sum-free Subset Theorem. arXiv preprint arXiv:1606.07823 (2016).

4 Problem 4

4.1 Question

Suppose $p > n > 10m^2$, with p prime, and let $0 < a_1 < a_2 < ... < a_m < p$ be integers. Prove that there is an integer x, $0 < x < p$ for which the m numbers

$$(xa_i \bmod p) \bmod n, 1 \leq i \leq m$$

are pairwise distinct.

4.2 Answer

Choose integer x uniformly and randomly in $[p-1]$, clearly, with the assumptions, for each i, $(xa_i \bmod p)$ is distinct in $[p-1]$. Let X be the random variable denoting the number of ordered pairs (i, j), where $i, j \in [m]$ and $i \neq j$, which is naturally the sum of indicator random variables X_{ij} for all possible (i, j), where X_{ij} indicates $(xa_i \bmod p) \bmod n = (xa_j \bmod p) \bmod n$. Thus,

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i\neq j} X_{ij} \\ &\leq \sum_i \sum_{j\neq i} \frac{\lfloor \frac{p-1}{n} \rfloor}{p-1} \\ &\leq \sum_i \sum_{j\neq i} \frac{1}{n} \\ &= \frac{m(m-1)}{n} \\ &< \frac{m(m-1)}{10m^2} \\ &< \frac{1}{10},\end{aligned}$$

which implies for some x, $X = 0$, i.e., $(xa_i \bmod p) \bmod n, 1 \leq i \leq m$ are pairwise distinct, completing the proof.

5 Problem 5

5.1 Question

Let H be a graph, and let $n > |V(H)|$ be an integer. Suppose there is a graph on n vertices and t edges containing no copy of H, and suppose that $tk > n^2 \log_e n$. Show that there is a coloring of the edges of the complete graph on n vertices by k colors with no monochromatic copy of H.

5.2 Answer

When $|V(H)| \leq 1$, it's not so meaningful, so we only need to care about the cases when $|V(H)| \geq 2$, and then $n \geq |V(H)| + 1 \geq 3$.
Let G be such a graph on n vertices and t edges containing no copy of H. As we supposed that $tk > n^2 \log_e n > \binom{n}{2}$, maybe we can find some way to "paste" k monochromatic copies of G onto K_n (some edges may be colored more than once, among the colors they are colored in, choosing any one can make the conclusion hold), where each edge is colored and no copy of H exists in the graph.
So, let's randomly "paste" k monochromatic copies of G onto K_n, mathematically speaking, let $V(G) = [n]$ and $V(K_n) = \{v_i\}_{i\in[n]}$, for each color, we randomly pick a permutation of $[n]$, $\sigma : [n] \to \sigma([n])$, then for all $(i, j) \in E(V(G))$, we color the edge $(v_{\sigma(i)}, v_{\sigma(j)})$ in K_n in that color (if it's already colored, then we simply recolor it). Let X be the random variable of edges not colored, which is naturally the summation of X_e, the indicator random variable of

e being not colored, for all $e \in E(K_n)$, thus clearly,

$$\mathbb{E}[X_e] = (1 - \frac{2t(n-2)!}{n!})^k = (1 - \frac{2t}{n(n-1)})^k.$$

Therefore,

$$\begin{aligned}
\mathbb{E}[X] =& \sum_e \mathbb{E}[X_e] \\
=& \sum_e (1 - \frac{2t}{n(n-1)})^k \\
=& \binom{n}{2} (1 - \frac{2t}{n(n-1)})^k \\
\leq& \frac{n(n-1)}{2} e^{-\frac{2tk}{n(n-1)}} \\
<& \frac{n^2}{2} e^{-2\log_e n} \\
=& \frac{1}{2},
\end{aligned}$$

as X is an integer, there is a coloring by pasting k monochromatic copies of G with $X = 0$, i.e., each edge is colored, completing the proof.

6 Problem 6

6.1 Question

(*) Prove, using the technique shown in The Probabilistic Lens: Hamiltonian Paths, that there is a constant $c > 0$ such that for every even $n \geq 4$ the following holds: For every undirected complete graph K on n vertices whose edges are colored red and blue, the number of alternating Hamiltonian cycles in K (i.e., properly edge-colored cycles of length n) is at most

$$n^c \frac{n!}{2^n}.$$

6.2 Answer

Let $n = 2k$, where integer $k \geq 2$. Let K be an undirected complete graph on n vertices whose edges are colored red and blue with equal probability $1/2$ uniformly and randomly. Construct a balanced directed complete bipartite graph G on $V(K)$ as follows, say L and R are the two disjoint parts of vertices in bipartite graph G, for any $v_1 \in L$ and $v_2 \in R$, if the edge between v_1 and v_2 is colored red in K, then in G, the edge between v_1 and v_2 is from v_1 to v_2, otherwise, it's from v_2 to v_1. Clearly, now the vertices of any directed Hamiltonian cycle in G originally form an alternating Hamiltonian cycle in K, and for each alternating Hamiltonian cycle A in G, let X_A denote the event that the vertices of A are alternatively assigned to L and R in G in any order of a traversal of A, i.e., these vertices form a directed Hamiltonian cycle in G, then we have

$$\Pr[X_A] = \frac{2}{\binom{n}{k}}.$$

Let $A(K)$ denote the number of alternating Hamiltonian cycles in K and let X be the random variable denoting the number of directed Hamiltonian cycles in G, we have

$$\mathbb{E}[X] = \frac{2A(K)}{\binom{n}{k}}.$$

On the other hand, for any possible G, let A_G be the adjacency matrix of G, we have

$$X \leq \mathrm{per}(A_G) \leq \prod_{i=1}^{n} (r_i!)^{1/r_i},$$

where r_i is the number of ones in the ith row of A_G, i.e., the out degree of the ith vertex in G. As we have $\sum_{i=1}^{n} r_i = k^2$, and observe the Lagrangian we can know $\prod_{i=1}^{n}(r_i!)^{1/r_i}$ is maximized when r_i is identical to each other for all $1 \leq i \leq n$, and $f(x) = (x!)^{1/x}$ is monotonic increasing on $\mathbb{Z}^+$, thus there is some G such that

$$\frac{2A(K)}{\binom{n}{k}} \leq X \leq \prod_{i=1}^{n} (r_i!)^{1/r_i} \leq (\lceil \frac{k}{2} \rceil !)^{\frac{n}{\lceil \frac{k}{2} \rceil}} \leq (\lceil \frac{k}{2} \rceil !)^4,$$

which implies that

$$\begin{aligned} 2A(K) \leq & \binom{n}{k} (\lceil \frac{k}{2} \rceil !)^4 \\ \leq & \frac{n!}{k!k!} (\lceil \frac{k}{2} \rceil !)^4 \\ \leq & \frac{n! n^5}{2^n}, \end{aligned}$$

where for the last inequality we used Stirling's approximation. Finally, we let $c = 5$ and complete the proof.

7 Problem 7

7.1 Question

Let $\mathcal{F}$ be a family of subsets of $N = \{1, 2, ..., n\}$, and suppose there are no $A, B \in \mathcal{F}$ satisfying $A \subset B$. Let $\sigma \in S_n$ be a random permutation of the elements of N and consider the random variable

$$X = |\{i : \{\sigma(1), \sigma(2), ..., \sigma(i)\} \in \mathcal{F}\}|.$$

By considering the expectation of X prove that $|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$.

7.2 Answer

Let N_i denote the number of elements (each element in $\mathcal{F}$ is a subset of N) in $\mathcal{F}$ with size i. Clearly $|\mathcal{F}| = \sum_{i=1}^{n} N_i$. By the definition of X and the assumption that there are no $A, B \in \mathcal{F}$ satisfying $A \subset B$, we can easily get $\mathbb{E}[X] \leq 1$ as $\{\sigma(1), \sigma(2), ..., \sigma(i)\} \subset \{\sigma(1), \sigma(2), ..., \sigma(j)\}$ for any $i < j$. Meanwhile we have

$$\mathbb{E}[X] = \sum_{i=1}^{n} \frac{N_i}{\binom{n}{i}},$$

thus,

$$\frac{|\mathcal{F}|}{\binom{n}{\lfloor n/2 \rfloor}} = \frac{\sum_{i=1}^{n} N_i}{\binom{n}{\lfloor n/2 \rfloor}} \leq \sum_{i=1}^{n} \frac{N_i}{\binom{n}{i}} \leq 1,$$

completing the proof.

7.3 A Failed Attempt

Inspired by problem 8 in chapter 1 (Prefix-free codes; Kraft Inequality), we can have do some inference as follows.

Let N_i denote the number of elements (each element in $\mathcal{F}$ is a subset of N) in $\mathcal{F}$ with size i. Consider the construction of the sets in $\mathcal{F}$ in the increasing order of size. When it comes to the sets with size j, each set should not contain any existing sets with size $< j$ as its subset, i.e., among all the $\binom{n}{j}$ trivial choices, each previous set with size k would forbid $\binom{n-k}{j-k}$ choices, besides, as we assume no member of $\mathcal{F}$ is a subset of another one, these forbidden sets are disjoint (WRONG!!!). Thus, for sets with size j, the size of feasible set is

$$\binom{n}{j} - \sum_{i=1}^{j-1} N_i \binom{n-i}{j-i},$$

which should satisfy

$$\binom{n}{j} - \sum_{i=1}^{j-1} N_i \binom{n-i}{j-i} \geq N_j.$$

We can rewrite it as

$$\binom{n}{j} \geq \sum_{i=1}^{j} N_i \binom{n-i}{j-i},$$

or

$$1 \geq \sum_{i=1}^{j} \frac{N_i \binom{j}{i}}{\binom{n}{i}}.$$

which holds for any $1 \leq j \leq n$.

8 Problem 8

8.1 Question

(*) Let X be a collection of pairwise orthogonal unit vectors in $\mathbb{R}^n$ and suppose the projection of each of these vectors on the first k coordinates is of Euclidean norm at least ϵ. Show that $|X| \leq k/\epsilon^2$, and this is tight for all $\epsilon^2 = k/2^r < 1$.

8.2 Answer

Let M be a matrix where each column vector M_i is the ith element of X. Let $|X| = p \leq n$, then $M \in \mathbb{R}^{n \times p}$, actually, M can be seen on Stiefel manifold $St(n,p)$ embedded in $\mathbb{R}^{n\times p}$, and we have $M^T M = I_p$. When $p = n$, $M \in O(n)$, the orthogonal group in dimension n, and we have $MM^T = M^T M = I_n$, otherwise, we can always find $M_\perp \in \mathbb{R}^{n\times(n-p)}$ such that

$\tilde{M} = [MM_\perp]$ (the concatenation of columns of M and $M_\perp$) is in $O(n)$. We will just use $\tilde{M}$ in the proof (for $p = n$, let $\tilde{M} = M$). The assumption means $\sum_{i=1}^{k} \tilde{M}_{is}^2 = \sum_{j=1}^{k} \tilde{M}_{sj}^2 \geq \epsilon^2$ for any $s \in [n]$, thus

$$p\epsilon^2 \leq \sum_{i=1}^{p} \sum_{j=1}^{k} \tilde{M}_{ij}^2 = \sum_{i=1}^{k} \sum_{j=1}^{p} \tilde{M}_{ij}^2 \leq \sum_{i=1}^{k} \sum_{j=1}^{n} \tilde{M}_{ij}^2 = k,$$

completing the proof for $|X| \leq k/\epsilon^2$.
When $\epsilon^2 = k/2^r < 1$, we want to find an example where $|X| = k/\epsilon^2 = 2^r \leq n$ to prove the tightness. Let $X = \{x_1, ..., x_{2^r}\}$, and we construct each $x_i \in X$ as follows. Let base vector $b = (2^{-r/2}, ..., 2^{-r/2}, 0, ..., 0)$ where the first 2^r coordinates are all $2^{-r/2}$ and the others are 0, and H be the Hadamard matrix of order 2^r. For each $i \in [n]$, let $x_i = (x_{ij})_{j=1}^{n}$, where

$$x_{ij} = \begin{cases} b_j H_{ij}, & j \leq 2^r \\ 0, & \text{otherwise} \end{cases}$$

We can easily check $|x_i| = |b| = 1$ for all $i \in [n]$, $x_i \cdot x_j = 0$ for all $i \neq j$ (with the property of Hadamard matrix), and $\sum_{l=1}^{k} x_{il}^2 = k/2^r = \epsilon^2$ (as $k < 2^r$), completing the proof.

9 Problem 9

9.1 Question

Let $G = (V, E)$ be a bipartite graph with n vertices and a list $S(v)$ of more than $\log_2 n$ colors associated with each vertex $v \in V$. Prove that there is a proper coloring of G assigning to each vertex v a color from its list $S(v)$.

9.2 Answer

Let $S = \bigcup_{v \in V} S(v)$, L and R denote the two parts of G, and let K be an array of length $|S|$ consisting of only 0s and 1s, where for any $1 \leq i \leq |S|$, $K_i = 0$ means color S_i should only be used in L, otherwise color S_i should only be used in R. If we can find such K where for each $v \in L$, we can always find some $1 \leq i \leq |S|$ such that $S_i \in S(v)$ and $K_i = 0$, meanwhile for each $v \in R$, we can always find some $1 \leq i \leq |S|$ such that $S_i \in S(v)$ and $K_i = 1$, then clearly there is a proper coloring.
Let X be the random variable denoting the number of vertices for which what we want cannot be satisfied and we assign the value of each K_i to be 0 or 1 with equal probability $1/2$, for each $v \in V$, the probability of what we want cannot be satisfied

$$\Pr[\text{fail}] < (\frac{1}{2})^{\log_2 n} = \frac{1}{n},$$

which implies that

$$\mathbb{E}[X] < \sum_{v \in V} \frac{1}{n} = 1.$$

As X is an integer, we can conclude that there is some K where $X = 0$, i.e., want we want is satisfied for all vertices, completing the proof.