CSS.205.1: Toolkit for TCS	Jan-May 2025
HW2 : MWUM and Linear Algebraic Method	
Out: 7 Mar, 2025	**Due:** 21 Mar, 2025

- *This homework has problems worth 30 points.*
- *Please take time to write clear and concise solutions. You are* STRONGLY *encouraged to submit LaTeXed solutions.*
- *Collaboration is OK, but please write your answers yourself, and include in your answers the names of* EVERYONE *you collaborated with and* ALL *references other than class notes you consulted. However,* ***not*** *acknowledging your collaborators and references will be treated as a serious case of* ***academic dishonesty****.*

1. **[From Satish Rao's homework assignments]** (8 points)

   In the experts framework, we compared an online algorithm that was allowed to pick among n (expert) strategies, with the best choice of a single expert in hindsight. In this question we first introduce a different measure of performance of the online algorithm, called *regret*. The regret is how much worse the online player does than the best offline player. i.e. if the best expert suffers a loss of L while the expected loss of the online player is $\mathbf{E}[L(A)]$, then the regret is $\mathbf{E}[L(A)] - L$.

   Assume that the game lasts for T rounds, and the loss in each round is in the interval $[0, 1]$.

   (a) (4 points) Show that the regret suffered by the experts algorithm (with appropriate parameters) compared to the best expert is $O\left(\sqrt{T \cdot \log n}\right)$.

   (b) (4 points) How important was it that the online player was allowed to switch between experts, while the offline player had to stick to a single expert? To answer this, let us consider the regret suffered by the online player compared to an offline player who is allowed to switch between experts at every step. Construct an example where L^*, the loss suffered by the offline player is 0, while the expected regret of A is at least $T \cdot (1 - 1/n)$.

2. **[Lower bound of 2 for any deterministic algorithm]** (5 points)

   In the lectures, we showed that the *weighted-majority (WM_η)* algorithm satisfies the following bound. For any positive integer T, let $m_i^{(T)}$ be the number of mistakes made by expert i upto step T and $M^{(T)}$ be the number of mistakes made by WM_η algorithm with parameter $\eta \in (0, 1/2]$. Then, for any expert $i \in [n]$, we have

   $$M^{(T)} \leq 2(1+\eta) \cdot m_i^{(T)} + \frac{2 \ln n}{\eta}.$$

   In other words, the WM_η algorithm makes no more than approximately *twice* the number of mistakes made by any expert. Show that no *deterministic* online algorithm can guarantee a perfomance better than twice the number of mistakes made by the best expert.

3. **[ε-Nets & Hitting Sets (modified from Anupam Gupta and Ryan O'Donnell's psets)]** (7 points)

   Let $(U, \mathcal{F})$ be a set system, where the universe U is of size n (i.e, $|U| = n$), and $\mathcal{F}$ is a collection of m sets $\{S_1, S_2, \ldots, S_m\}$ where each $S_i \subseteq U$. Define the following two quantities:

   Hitting Set: The set $H \subseteq U$ is a hitting set for $\mathcal{F}$ if $H \cap S_i \neq \emptyset$ for all $i \in [m]$. Let c be the size of the smallest hitting set for $\mathcal{F}$; note that this is NP-hard to compute.

   ε-net: Given non-negative weights w_e for each element $e \in U$, define the weight of a set A as $w(A) = \sum_{e \in A} w_e$. A set $N \subseteq U$ is an ε-net for $(\mathcal{F}, w)$ if for all sets S_i such that $w(S_i) \geq \varepsilon \cdot w(U)$, we have $N \cap S_i \neq \emptyset$. (In other words, N hits all the "heavy-weight" sets.)

   (a) (2 points) For any parameter $\varepsilon > 0$, give a polynomial-time algorithm that given any set system $(U, \mathcal{F})$ with m sets and a weight function $w : U \to \mathbb{R}_{\geq 0}$ finds an ε-net of size at most $O(\log m / \varepsilon)$.

   Hint: Check if a greedy algorithm (or even a randomized algorithm) works.

(b) (4 points) In this part, we will use the above algorithm for ε-nets to construct a hitting set that is not too large compared the size of the optimal hitting set.

Suppose you are given an algorithm E_ε that when given a set family $(U, \mathcal{F})$ and an associated weight function w, finds an ε-net of size $T(\varepsilon, m)$. (This could be, for instance the algorithm designed in part 3a).

Consider the following algorithm that uses E_ε to compute a hitting set for $\mathcal{F}$:

Data: $(U, \mathcal{F})$ - a set system
Result: A hitting set H
Set $w(e) \leftarrow 1$ for all $e \in U$;
repeat
  Use algorithm E_ε to find an ε-net H for $(\mathcal{F}, w)$;
  Let S be any set in $\mathcal{F}$ such that $S \cap H = \emptyset$. If no such set exists, set $S \leftarrow \emptyset$;
  for *all elements* $e \in S$ **do**
    $w(e) \leftarrow 2 \cdot w(e)$ (i.e., double its weight);
  end
until H *is a hitting set for* $\mathcal{F}$;
Ouput H;

Algorithm 1: Hitting Set Algorithm From ε-nets

Let c be the size of an optimal hitting set S^* of $(U, \mathcal{F})$ (note that it is NP-hard to determine this quantity). Show that if $\varepsilon < 2^{1/c} - 1$, then the above algorithm terminates within

$$\frac{\log_2\left(\frac{n}{c}\right)}{\frac{1}{c} - \log_2(1+\varepsilon)}$$

iterations of the return loop at which point it outputs a hitting set of size at most $T(\varepsilon, m)$.

One such choice of ε is $1/2c$, which gives a bound of $O(c \cdot \log(n/c))$ on the number of iterations.

> **Hint:** Compare the weight $w(U)$ of the universe U and the weight $w(S^*)$ of the optimal hitting set S^* at the end of each iteration of the return loop.

(c) (1 point) The above algorithm requires knowledge of the optimal hitting set size c. Use a doubling argument or otherwise, to give an efficient $O(\log m)$-approximation algorithm of the size of the optimal hitting set size

4. **[3-AP-free sets in $\mathbb{F}_3^n$ via the polynomial method]** (10 points)

Let $A \subseteq \mathbb{F}_3^n$. We say that A is 3-AP-free if there does not exist $x \neq y \in \mathbb{F}_3^n$ such that $x, (x+y)/2, y \in A$ (i.e., A does not contain any non-trivial arithmetic progression of length 3). In this problem, we will use the polynomial method to show that any 3-AP-free set is of size at most c^n for some fixed $c \in (2, 3)$.

(a) (2 points) Let $0 \leq d \leq 2n$. Let $V_d(n)$ denote the set of all functions from $\mathbb{F}_3^n$ to $\mathbb{F}_3$ expressible as degree d polynomials. In other words, if $f \in V_d$, then f can be expressed as a polynomial of the form

$$f(x_1, \ldots, x_n) = \sum_{\mathbf{a}=(a_1,\ldots,a_n)\in\{0,1,2\}^n : \sum a_i \leq d} c_{\mathbf{a}} \prod_{i=1}^{n} x_i^{a_i}.$$

Let $m_d(n) = \dim(V_d(n))$. Prove the following facts about m_d.

i. $m_{2n}(n) = 3^n$.
ii. For all $0 \leq d \leq 2n$, $m_{2n-d}(n) = 3^n - m_d(n)$.
iii. There exists $c \in (2, 3)$ such that $m_{2n/3}(n) \leq c^n$.

(b) (3 points) Let $A \subseteq \mathbb{F}_3^n$ be 3-AP-free.

i. Show that if $m_d > 3^n - |A|$, then there exists a non-zero $f \in V_d$ such that $f(\mathbf{x}) = 0$ for all $\mathbf{x} \notin A$.
ii. Strengthen the above to show that if $m_d > 3^n - |A|$, then there exists an $f \in V_d$ such that $f(\mathbf{x}) = 0$ for all $\mathbf{x} \notin A$ and f is non-zero on at least $(m_d + |A| - 3^n)$ points in A.
iii. Let $f : \mathbb{F}_3^n \to \mathbb{F}_3$ such that $f(\mathbf{x}) = 0$ for all $\mathbf{x} \notin A$. Define the matrix $M_f \in \mathbb{F}_3^{A \times A}$ as follows: $M_f(x, y) := f((x+y)/2)$ for all $x, y \in A$. Show that the rank of M_f is exactly $|\{\mathbf{x} \in A | f(\mathbf{x}) \neq 0\}|$.

(c) (4 points) Let $g : \mathbb{F}_3^n \to \mathbb{F}_3$ be a function in $V_d(n)$. Consider the matrix M_g given by $M_g(x, y) := g(x+y)$. Prove the following facts about the rank of the matrix M_g.

i. $\text{rank}(M_g) \leq m_d(n)$.

ii. Strenghten the above to show that $\text{rank}(M_g) \leq 2 \cdot m_{d/2}(n)$.

Hint: Recall that if a $t \times t$-matrix M can be decomposed as $M = UV$ where U is a $t \times r$-matrix and V is a $r \times t$ matrix (or equivalently there exists $2t$ r-dimensional vectors $\mathbf{u}_1, \ldots, \mathbf{u}_t, \mathbf{v}_1, \ldots, \mathbf{v}_t$ such that $M(i, j) = u_i^T v_i$), then $\text{rank}(M) \leq r$.

(d) (1 point) Conclude from the above parts that if A is 3-AP-free, then $|A| \leq m_{2n-d} + 2m_{d/2}$. Setting $d = 4n/3$ show that $|A| \leq 3c^n$ where c is as in Part 4(a)iii