Homework #1
CS 6501: Learning and Game Theory (Fall'19)

Due Thursday 09/19 3:30 pm

General Instructions The assignment is meant to be challenging. Feel free to discuss with fellow students, however please write up your solutions independently (e.g., start writing solutions after a few hours of any discussion) and acknowledge everyone you discussed the homework with on your writeup. The course materials are all on the course website: http://www.haifeng-xu.com/cs6501fa19. You may refer to any materials covered in our class. However, any attempt to consult outside sources, on the Internet or otherwise, for solutions to any of these homework problems is *not* allowed.

Whenever a question asks you to "show" or "prove" a claim, please provide a formal mathematical proof. These problems have been labeled based on their difficulties. `Short` problems are intended to take you 5-15 minutes each and `medium` problems are intended to take 15-30 minutes each. `Long` problems may take anywhere between 30 minutes to several hours depending on whether inspiration strikes.

Finally, please write your solutions in latex — hand written solutions will not be accepted. Hope you enjoy the homework!

Problem 1

Consider a Linear Program (LP) in the following standard form where $c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

$$\begin{array}{ll} \text{maximize} & c^T \cdot x \\ \text{subject to} & Ax \leq b \\ & x \geq 0 \end{array} \tag{1}$$

Prove the following facts about the LP.

1. (`Short,3 points`) The feasible region of the LP is a convex set.
2. (`Short,3 points`) At any vertex of the feasible region, n linearly independent constraints are satisfied with equality (a.k.a. *tight*).
3. (`Short,3 points`) The set of optimal solutions of the LP is a convex set.
4. (`Short,3 points`) We learned in class that the dual of LP (1) is the following LP (2)

$$\begin{array}{ll} \text{minimize} & b^T \cdot y \\ \text{subject to} & A^T y \geq c \\ & y \geq 0 \end{array} \tag{2}$$

Prove that the dual of LP (2) is the original LP (1).

Problem 2

Prove the following projection lemma and separating hyperplane theorem.

1. (**Projection Lemma**, `Medium, 5 points`) Let $Z \subset \mathbb{R}^n$ be a nonempty closed convex set and $y \notin Z$ be any point in $\mathbb{R}^n$. Prove that there exists $z^* \in Z$ that has the minimum l_2 distance from y among all $z \in Z$. Moreover, $\forall z \in Z$ we have $(y - z^*)^T \cdot (z - z^*) \leq 0$. (hint: use Weierstrass' Theorem).

2. (**Separating Hyperplane Theorem**, `Short, 3 points`) Let $Z \subset \mathbb{R}^n$ be a nonempty closed convex set and let $y \notin Z$ be any point in $\mathbb{R}^n$. Prove that there exists a hyperplane $\alpha^T \cdot x = \beta$ that strictly separates y from Z. That is, $\alpha^T \cdot z \geq \beta$ for any $z \in Z$ but $\alpha^T \cdot y < \beta$.

Problem 3: Linear Programming for Machine Learning

In this question, you will learn to formulate some machine learning problems as linear programs. Let us assume that there are n data points $(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_n, y_n)$ where $\mathbf{x}_i \in \mathbb{R}^m$ is interpreted as an m-dimensional feature vector and $y_i \in \mathbb{R}$ is the corresponding label.

1. (**Data Fitting**, `Short, 4 points`) We want to construct a linear predictive model $\mathbf{a} \cdot \mathbf{x}$ to fit the value y. One possible loss function of this fitting is the worst-case error, i.e., $\max_i |\mathbf{a} \cdot \mathbf{x}_i - y_i|$. Show that computing the linear predictive model that minimizes the worst-case error can be formulated as a linear program.

2. (**Linear Classification**, `Short, 4 points`) When $y_i \in \{-1, 1\}$ is a binary label for data point i, this gives rise to a binary classification problem. In linear classification, we seek to find a hyperplane $\mathbf{a} \cdot \mathbf{x} - b = 0$ that strictly separates the data points with label 1 from the points with label -1. That is, $\mathbf{a} \cdot \mathbf{x}_i - b > 0$ if $y_i = 1$ and $\mathbf{a} \cdot \mathbf{x}_i - b < 0$ if $y_i = -1$ (note that the requirement "$<$" or "$>$" is strict). For convenience, we also call such a hyperplane *separating hyperplane*. Show that computing a separating hyperplane or asserting that it does not exist can be formulated as a linear feasibility problem (i.e,, a linear program with an arbitrary objective function).

   (Note that in linear programs, any linear constraint *cannot* have strict inequalities like "$>$" or "$<$"; see the general form of LPs in Lecture 2 slides.)

Problem 4: Rock-Paper-Scissor

In this problem, you will learn to master the rock-paper-scissor game. Recall that the game has the following payoff structure where each utility (x, y) means the row player receives x and the column player receives y.

	Rock	Paper	Scissor
Rock	(0, 0)	(-1, 1)	(1, -1)
Paper	(1, -1)	(0, 0)	(-1, 1)
Scissor	(-1, 1)	(1, -1)	(0, 0)

Table 1: Payoffs of the Standard Rock-Paper-Scissor Game

1. (`Short, 3 points`) Prove that the above rock-paper-scissor has a *unique* Nash equilibrium, which is that each player picks one of $\{Rock, Paper, Scissor\}$ uniformly at random.

2. (`Medium, 5 points`) Consider the situation where the column player is forbidden to play $Scissor$ (equivalently, the last column of the above payoff matrix is deleted). What is the Nash equilibrium of this new variant of the game.

3. (`Short, 3 points`) Consider the situation where the two players are encouraged to collaborate. In particular, if they play the same action, each will receive 0.5. This results in the following game variant. What is the Nash equilibrium of this new game?

	Rock	Paper	Scissor
Rock	(0.5, 0.5)	(-1, 1)	(1, -1)
Paper	(1, -1)	(0.5, 0.5)	(-1, 1)
Scissor	(-1, 1)	(1, -1)	(0.5, 0.5)

Table 2: Payoffs of the Rock-Paper-Scissor Game with Encouraged Collaboration

4. (`Medium, 5 points`) Imagine that you are an outsider who watches two players playing the above game variant with encouraged collaboration, and you can recommend actions to the two players using a correlated equilibrium. If you want to *maximize* the sum of the two players' expected utilities, which correlated equilibrium should you use? If you want to *minimize* the sum of their expected utilities, which correlated equilibrium should you use?

Problem 5: Stackelberg Games

In this problem, you will learn another type of games called **Stackelberg games**. A Stackelberg game is a two-player game but with *sequential* player moves. In particular, a normal-form Stackelberg game is described by two matrices $A, B \in \mathbb{R}^{n \times m}$ where A is the payoff matrix of the row player who has action set $[n] = \{1, \cdots, n\}$ and B is the payoff matrix of the column player who has action set $[m] = \{1, \cdots, m\}$. The row player moves first (call *her* the *leader*) and the column player (call *him* the *follower*) moves second and thus can see the row player's strategy and then responds with his best action. Similar to the argument we saw in class, such a best response can without loss of generality be a pure best response. Sometimes there may be multiple best responses. In this case we assume that the follower is a benign player so that he will always pick the one that is the best for the leader, i.e., the follower breaks ties in favor of the leader.

It is not difficult to see that, after seeing the leader's strategy — either pure strategy or mixed strategy — the follower's best response action is easy to compute. That is, simply check the utility of each follower action $j \in [m]$ and then pick the best one. Therefore, research in Stackelberg games mainly focuses on computing the optimal leader strategy, which is also called the leader's Strong Stackelberg Equilibrium (SSE) strategy.

Answer the following questions about Stackelberg games.

1. (`Short, 3 points`) **A warm-up example.** Recall the traffic light game from Lecture 4, as follows. Assume that the row player is the leader and she can only play a pure strategy[1], what is the leader's SSE strategy?

[1] For example, maybe because the follower can observe whatever pure action the leader takes.

	STOP	GO
STOP	(-3, -2)	(-3, 0)
GO	(0, -2)	(-100, -100)

Table 3: Payoffs of the Traffic Light Game

2. (`Short,3 points`) Consider the normal-form Stackelberg game and assume that the leader can only play a pure strategy. Show that there is a $\mathcal{O}(nm)$ time algorithm that computes the leader's pure SSE strategy.

3. (`Medium,5 points`) We now consider the case where the leader can play a mixed strategy. To compute the leader's SSE (mixed) strategy, consider the following simpler *SSE with promise* problem. That is, imagine that there is an oracle who promises us that when the leader plays the mixed SSE strategy, the follower's best response action will be j^*. Show that given this credible promise, the leader's SSE strategy can be computed by a linear program.

   Use one or two sentences to briefly explain how we can still compute the leader's SSE strategy efficiently even without the oracle's promise, by solving m linear programs.

4. (`Medium,5 points`) Prove that the leader's utility by playing the SSE mixed strategy (and the follower will best respond) is at least her utility in any Nash equilibrium of the game when players move simultaneously.

Problem 6: Boosting (`Long,10 points`)

A fundamental concept in learning theory is *boosting*, intuitively means that classifiers that perform only slightly better than random guess can be turned into a classifier that is never wrong. In this question, you will prove a basic version of this celebrated result using the minimax theorem for zero-sum games.

Let $\mathcal{X} = \{x_1, \cdots, x_n\}$ be any feature space and $\mathcal{H} = \{h : X \to \{-1, 1\}\}$ be a set of classifiers over $\mathcal{X}$ (a.k.a., hypothesis class). For example, $\mathcal{H}$ could be the set of all linear classifiers. However, for simplicity, in this question we will assume that $\mathcal{H} = \{h_1, \cdots, h_m\}$ is finite. Let $g : \mathcal{X} \to \{-1, 1\}$ be the ground truth, i.e., the true label of x_j is $g(x_j)$.

The *weak learnability assumption* on $\mathcal{H}$ says that $\mathcal{H}$ is good in the following sense: there exists $\epsilon > 0$ such that for any distribution $p(\in \Delta_n)$ over $\mathcal{X}$, there exists a classifier h_i such that h_i is correct with probability at least $\frac{1}{2} + \epsilon$ for point x drawn from p, or more formally,

$$\sum_{j=1}^{n} p_j \cdot \mathbb{I}[h_i(x_j) = g(x_j)] \geq \frac{1}{2} + \epsilon,$$

where $\mathbb{I}[h_i(x_j) = g(x_j)]$ is the indicator function. That is, $\mathbb{I}[h_i(x_j) = g(x_j)]$ equals 1 if $h_i(x_j) = g(x_j)$ and equals 0 otherwise.

It turns out that weak learnability implies something much stronger — we can combine classifiers in $\mathcal{H}$ to construct a classifier that is always correct (a.k.a., *strong learnability*), formally stated as follows.

If $\mathcal{H}$ satisfies the weak learnability assumption, then there always exists a distribution $q(\in \Delta_m)$ over $\mathcal{H}$ such that the following weighted classifier:

$$h_q(x) = \begin{cases} 1 & if \quad \sum_{i=1}^{n} q_i h_i(x) \geq 0 \\ 0 & otherwise \end{cases}$$

is always correct, that is, $h_q(x) = g(x)$ *for any* $x \in \mathcal{X}$.

Prove the above statement.

[Hint: The classification problem can be viewed as a zero-sum game played between a *classifier designer* whose pure strategy is to pick a classifier from $\mathcal{H}$ and an *adversary* whose pure strategy is to pick a data point from $\mathcal{X}$. Think about how to define the payoff matrix of this game and what weak learnability means in the zero-sum game context.]

Problem 7: Optimal Strategic Attack to ML Algorithms (`Long, 10 points`)

Strategic or adversarial attacks to machine learning algorithms has been a hot research topic recently. In this question, you will device an *optimal strategic attack* to the machine learning algorithm discussed in Lecture 1 of our class. In particular, we studied the problem of selling a product (with unlimited supply) to N sequentially arriving buyers. All the buyers have the same value $v \in [0, 1]$ for the product but v is unknown to the seller. In order to maximize the seller's revenue, we described an online learning algorithm for selling the product that achieves regret $(2 \log \log N + 1)$.

In this question, we concern a slight variant of the above problem. That is, the seller sells the product (with unlimited supply) to a *single buyer* who repeatedly shows up for N rounds. The buyer's value $v \in [0, 1]$ is unknown to the seller. We assume that the seller still uses exactly the same algorithm as we described in class (you may need to review lecture 1 slides if necessary). Naturally, knowing that the seller is learning his value, the buyer will be strategic about his response at each round. For example, when offered a price p_n at round n, the buyer may intentionally respond with "Reject" even though $v > p_n$ because this will trick the seller to offer lower prices in next rounds. On the other hand, a "Reject" response also leads to buyer utility 0 whereas an "Accept" could have given him a utility of $v - p_n (> 0)$ at round n. Therefore, the strategic buyer who looks to *maximize his total utility* would need to balance between using "Reject" to induce lower prices and using "Accept" to collect positive utilities.

More formally, denote the seller's price at round n by $p_n \in [0, 1]$ and the buyer's response by $s_n \in \{0, 1\}$ for $n = 1, \cdots, N$, where $s_n = 1$ means the buyer responds with "Accept" at round n and $s_n = 0$ means a buyer response of "Reject". The total utility of the buyer is thus $\sum_{n=1}^{N} s_n(v - p_n)$. Assume that the seller is committed to run the algorithm as described in class (the one achieving $(2 \log \log N + 1)$ regret for N repeated buyers), what is the optimal buyer response strategy $s = (s_1, \cdots, s_N)$? Prove your answer. The solution should concretely specify the optimal buyer strategy in terms of the parameter N and v.

What if the seller runs the standard binary search algorithm? What is the optimal buyer response strategy? Will the buyer gain less or more utility in this case? Prove your answers.

Hint: think about the following question — if the buyer will reject the offer for k rounds for some $k \leq N$, which k of the N rounds should he choose to reject the offer so that it maximizes his utility?