TTIC 31250 : Introduction to Theory of Machine Learning

Spring 2020

Homework 5

Solutions by

Jinhong Du

dujinhong@uchicago.edu

12243476

Exercises

1. **Zero-sum Games.** Consider the following zero-sum game. Player A (Alice) hides either a nickel or a quarter behind her back. Then, player B (Bob) guesses which it is. If Bob guesses correctly, he wins the coin. If Bob guesses incorrectly, he has to pay Alice 15 cents. In other words, the amount that Alice wins can be summarized by the following payoff matrix:

Alice hides		Bob guesses N	Bob guesses Q
	N	-5	$+15$
	Q	$+15$	-25

This seems like a fair game since when Bob loses, he pays Alice the average of 5 and 25, but we will see that one of the players in fact has an advantage.

(a) What is the value to Alice of the strategy "with probability $\frac{1}{2}$ hide a nickel and with probability $\frac{1}{2}$ hide a quarter"? (The value of a strategy is its value assuming that the opponent knows it and plays a best-response to it).

Suppose that B guesses a nickel and a quarter with probability q and $1-q$ respectively. Then the payoff of B is

$$5 \times \frac{1}{2}q - 15 \times \frac{1}{2}(1-q) - 15 \times \frac{1}{2}q + 25 \times \frac{1}{2}(1-q) = 5(1-2q) \leq 5$$

with equality if and only if $q = 0$. So the best strategy for B is to always guess a quarter. Therefore, the value to A of this strategy is

$$15 \times \frac{1}{2} - 25 \times \frac{1}{2} = -5.$$

(b) What is Alice's minimax optimal strategy, and what is its value?

Proof. Suppose that A hides a nickel and a quarter with probability p and $1-p$ respectively, and B guesses a nickel and a quarter with probability q and $1-q$ respectively. Then the payoff of B is

$$\begin{aligned} &5pq - 15p(1-q) - 15(1-p)q + 25(1-p)(1-q) \\ =&5[(5-8p)(1-q) - (3-4p)q] \\ =&5[(5-8p) - (8-12p)q]. \end{aligned}$$

When $8-12p > 0$ i.e. $p < \frac{2}{3}$, $5[(5-8p)-(8-12p)q] \leq 5(5-8p)$ with equality if and only if $q = 0$.
When $8-12p < 0$ i.e. $p > \frac{2}{3}$, $5[(5-8p)-(8-12p)q] \leq 5(4p-3)$ with equality if and only if $q = 1$.
When $8-12p = 0$, i.e. $p = \frac{2}{3}$, it doesn't matter what value q is.
Since

$$\min_{p \leq \frac{2}{3}} 5(5-8p) = \min_{p \geq \frac{2}{3}} 5(4p-3) = -\frac{5}{3}$$

with minima attained when $p = \frac{2}{3}$, so A's minimax optimal strategy is to hides a nickel and a quarter with probability $\frac{2}{3}$ and $\frac{1}{3}$ respectively. As the payoff of B is $-\frac{5}{3}$ when A plays a minimax optimal strategy, the value of A is $\frac{5}{3}$. □

(c) What is Bob's minimax optimal strategy, and what is its value to Bob?

Suppose that A hides a nickel and a quarter with probability p and $1-p$ respectively, and B guesses a nickel and a quarter with probability q and $1-q$ respectively. Then the payoff of A is

$$-5pq + 15p(1-q) + 15(1-p)q - 25(1-p)(1-q) = 5[(8q-5) + (8-12q)p].$$

When $8-12q > 0$ i.e. $q < \frac{2}{3}$, $5[(8q-5)+(8-12q)p] \le 5(3-4q)$ with equality if and only if $p = 1$.
When $8-12q < 0$ i.e. $q > \frac{2}{3}$, $5[(8q-5)+(8-12q)p] \le 5(8q-5)$ with equality if and only if $p = 0$.
When $8-12q = 0$, i.e. $q = \frac{2}{3}$, it doesn't matter what value p is.
Since

$$\min_{q \le \frac{2}{3}} 5(3-4q) = \min_{q \ge \frac{2}{3}} 5(8q-5) = \frac{5}{3}$$

with minima attained when $q = \frac{2}{3}$, so B's minimax optimal strategy is to guess a nickel and a quarter with probability $\frac{2}{3}$ and $\frac{1}{3}$ respectively. As the payoff of A is $\frac{5}{3}$ when B plays a minimax optimal strategy, the value of B is $-\frac{5}{3}$.

(d) Is it better to be Alice or Bob in this game?

Proof. As the values for A and B are $\frac{5}{3}$ and $-\frac{5}{3}$, when they play their own minimax optimal strategies, it is better to be Alice in this game. □

Problems

2. **On approximate Nash equilibria.** Consider a two-player general-sum game. Let us for concreteness focus on games where each player has n actions, and use R to denote the payoff matrix for the row player and C to denote the payoff matrix for the column player. (So if the row-player plays action i and the column-player plays action j, then the row-player gets R_{ij} and the column-player gets C_{ij}. Recall that a Nash Equilibrium is a pair of distributions p and q (one for each player) such that neither player has any incentive to deviate from its distribution assuming that the other player doesn't deviate from its distribution either. Formally, a pair of distributions p (for the row player) and q (for the column player) is a Nash equilibrium if the following holds: assuming the column player plays at random from q, the expected payoff to the row player for each row i with $p_i > 0$ is equal to the maximum payoff out of all the rows ($e_i^\top Rq = \max_{i'} e_{i'}^\top Rq$); and, assuming the row player plays at random from p, the expected payoff to the column player for each column j with $q_j > 0$ is equal to the maximum payoff out of all the columns ($p^\top Ce_j = \max_{j'} p^\top Ce_{j'}$). (Here, e_i denotes the column-vector with a 1 in position i and 0 everywhere else).

   Now, assume we have a game in which all payoffs are in the range $[0, 1]$. Define a pair of distributions p, q to be an "ϵ-Nash" equilibrium if each player has *at most* ϵ incentive to deviate. That is, the expected payoff to the row player for each row i with $p_i > 0$ is within ϵ of the maximum payoff out of all the rows, and vice-versa for the column player.

   Using the fact that Nash equilibria must exist, show that there must exist an ϵ-Nash equilibrium in which each player has positive probability on at most $O(\frac{1}{\epsilon^2} \log n)$ actions (rows or columns).

   Hint #1: what is a good randomized way to get a sparse approximation to a probability distribution p that was handed to you?

   Hint #2: your solution will require using Hoeffding bounds and the union bound.

   Note: this fact yields an $n^{O(\frac{1}{\epsilon^2} \log n)}$-time algorithm for finding an ϵ-Nash equilibrium. No PTAS (algorithm running in time polynomial in n for any fixed $\epsilon > 0$) is known, however.

   Proof. We independently sample k times from p and q to form multisets (which allow repeated elements) of actions $A_{R,k}$ and $A_{C,k}$, respectively. Then defind strategies $a_k, b_k \in \mathbb{R}^n$ by uniformly sampling from $A_{R,k}$ and $A_{C,k}$, and set $a_{k,j}$ (or $b_{k,j}$) to be 0 if action j is not in $A_{R,k}$ (or $A_{C,k}$).
   Notice that when the strategies of the row and column players are p and q, the expected payoffs to them are $p^\top Rq$ and $p^\top Cq$, respectively. We define the following events:

$$
\begin{aligned}
\phi_R &= \{|p^\top Rq - a_k^\top Rb_k| < \frac{\epsilon}{2}\} \\
\phi_C &= \{|p^\top Cq - a_k^\top Cb_k| < \frac{\epsilon}{2}\} \\
\pi_{R,j} &= \{e_j^\top Rb_k < a_k^\top Rb_k + \epsilon\} \\
\pi_{C,j} &= \{a_k^\top Ce_j < a_k^\top Cb_k + \epsilon\} \\
E_k &= \phi_R \cap \phi_C \bigcap_{j=1}^{n} (\pi_{R_j} \cap \pi_{C_j}).
\end{aligned}
$$

   Then a_k and b_k are Nash Equilibrium if and only if $\bigcap_{j=1}^n (\pi_{R_j} \cap \pi_{C_j})$ happens. So we just need to show that for some k, $\mathbb{P}(E_k) > 0$.

Solution (cont.)

Let

$$\phi_{R1} = \{|p^\top Rq - a_k^\top Rq| < \frac{\epsilon}{4}\}$$
$$\phi_{R2} = \{|a_k^\top Rq - a_k^\top Rb_k| < \frac{\epsilon}{4}\}$$
$$\phi_{C1} = \{|p^\top Cq - p^\top Cb_k| < \frac{\epsilon}{4}\}$$
$$\phi_{C2} = \{|p^\top Cb_k - a_k^\top Cb_k| < \frac{\epsilon}{4}\}.$$

Since

$$|p^\top Rq - a_k^\top Rb_k| \le |p^\top Rq - a_k^\top Rq| + |a_k^\top Rq - a_k^\top Rb_k|$$
$$|p^\top Cq - a_k^\top Cb_k| \le |p^\top Cq - p^\top Cb_k| + |p^\top Cb_k - a_k^\top Cb_k|,$$

we have $\phi_{R1} \cap \phi_{R2} \subseteq \phi_R$ and $\phi_{C1} \cap \phi_{C2} \subseteq \phi_C$. So $\phi_R^C \subseteq \phi_{R1}^C \cup \phi_{R2}^C$ and $\phi_C^C \subseteq \phi_{C1}^C \cup \phi_{C2}^C$. As $a_k^\top Rq$ is the sum of k indendent random variables each of expected value $p^\top Rq$, by Hoeffding bounds, we have $\mathbb{P}(\phi_{R1}^C) \le 2e^{-\frac{k\epsilon^2}{8}}$. Analogously, we have

$$\mathbb{P}(\phi_{R1}^C) \le 2e^{-\frac{k\epsilon^2}{8}}, \qquad \mathbb{P}(\phi_{R2}^C) \le 2e^{-\frac{k\epsilon^2}{8}}, \qquad \mathbb{P}(\phi_{C1}^C) \le 2e^{-\frac{k\epsilon^2}{8}}, \qquad \mathbb{P}(\phi_{C2}^C) \le 2e^{-\frac{k\epsilon^2}{8}}.$$

So

$$\mathbb{P}(\phi_R^C) \le \mathbb{P}(\phi_{R1}^C) + \mathbb{P}(\phi_{R2}^C) \le 4e^{-\frac{k\epsilon^2}{8}}$$
$$\mathbb{P}(\phi_C^C) \le \mathbb{P}(\phi_{C1}^C) + \mathbb{P}(\phi_{C2}^C) \le 4e^{-\frac{k\epsilon^2}{8}}.$$

Define

$$\psi_{Rj} = \{e_j^\top Rb_k < e_j^\top Rq + \frac{\epsilon}{2}\}$$
$$\psi_{Cj} = \{a_k^\top Re_j < p^\top Re_j + \frac{\epsilon}{2}\}.$$

As $\psi_{Rj} \cap \phi_R \subseteq \pi_{R,j}$ and $\psi_{Cj} \cap \phi_C \subseteq \pi_{C,j}$, we have

$$\mathbb{P}(\pi_{Rj}^C) \le \mathbb{P}(\psi_{Rj}^C) + \mathbb{P}(\phi_R^C) \le e^{-\frac{k\epsilon^2}{2}} + 4e^{-\frac{k\epsilon^2}{8}},$$
$$\mathbb{P}(\pi_{Cj}^C) \le \mathbb{P}(\psi_{Cj}^C) + \mathbb{P}(\phi_C^C) \le e^{-\frac{k\epsilon^2}{2}} + 4e^{-\frac{k\epsilon^2}{8}}.$$

So

$$\begin{aligned}\mathbb{P}(E_k^C) &\le \mathbb{P}(\phi_R^C) + \mathbb{P}(\phi_C^C) + \sum_{j=1}^n \mathbb{P}(\pi_{Rj}^C) + \sum_{j=1}^n \mathbb{P}(\pi_{Cj}^C) \\ &\le 8e^{-\frac{k\epsilon^2}{8}} + 2n(e^{-\frac{k\epsilon^2}{2}} + 4e^{-\frac{k\epsilon^2}{8}}) \\ &\lesssim ne^{-k\epsilon^2}\end{aligned}$$

Setting $ne^{-k\epsilon^2} \le 1$ yields $k \ge \frac{1}{\epsilon^2}\log n$. Therefore, if $k = O(\frac{1}{\epsilon^2}\log n)$, there exist ϵ-Nash Equilibrium strategies a_k and b_k for two players with positive probability on at most $O(\frac{1}{\epsilon^2}\log n)$ actions.

$\square$

3. **Compression bounds.** For some learning algorithms, the hypothesis produced by running the algorithm on a training set of size n can be uniquely described by giving k of the training examples. E.g., if you are learning an interval on the line using the simple algorithm "take the smallest interval that encloses all the positive examples," then the hypothesis can be reconstructed from just being told the outermost positive examples, so $k = 2$. For a conservative Mistake-Bound learning algorithm, you can reconstruct the hypothesis produced by the algorithm by just looking at the examples on which a mistake was made, so $k \leq M$, where M is the algorithm's mistake-bound. (In this case, you would also care about the order in which those examples arrived.)

   Your job in this problem is to prove a PAC generalization guarantee based on k (essentially, proving that if k is small, then this is a legitimate notion of a "simple" hypothesis; these are called *compression bounds*). Specifically, assume we fix a reconstruction procedure, so that for a given sequence of examples S' we have a well-defined hypothesis $h_{S'}$. You will show that

$$\Pr_{S \sim D^n}\left(\exists S' \subseteq S, |S'| = k, \text{ such that } h_{S'} \text{ has 0 error on } S - S' \text{ but true error } > \epsilon\right) \leq \delta$$

   so long as

$$n \geq \frac{1}{\epsilon}\left(k \ln n + \epsilon k + \ln \frac{1}{\delta}\right).$$

   (a) First, prove the following easier statement. Let's use $x_1, \dots, x_n$ to denote the examples in S. Now suppose you are given a sequence of indices $i_1, \dots, i_k$. Define $A_{i_1,\dots,i_k}$ to be the event that $h_{(x_{i_1},\dots,x_{i_k})}$ has zero error on all examples $x_j \in S$ such that $j \notin \{i_1, \dots, i_k\}$ and yet the true error of $h_{(x_{i_1},\dots,x_{i_k})}$ is more than ϵ. Prove that if $S \sim D^n$, the probability of event $A_{i_1,\dots,i_k}$ is at most $(1-\epsilon)^{n-k}$.

   Proof. Let $S = S_1 \times S_2$ where $S_1 \sim D^k$ and $S_2 \sim D^{n-k}$, and S_1 contains examples with indeces $i_1, \dots, i_k$. Since samples from S_1 and S_2 are independent, we have

$$\begin{aligned}
\mathbb{P}_{S \sim D^n}(A_{i_1,\dots,i_k}) &= \mathbb{P}_{S \sim D^n}(err_{S_1}(h_{S_1}) > \epsilon, err_{S_2}(h_{S_1}) = 0) \\
&= \mathbb{P}_{S \sim D^n}(err_{S_2}(h_{S_1}) = 0 | err_{S_1}(h_{S_1}) > \epsilon)\mathbb{P}_{S \sim D^n}(err_{S_1}(h_{S_1}) > \epsilon) \\
&\leq \mathbb{P}_{S \sim D^n}(err_{S_2}(h_{S_1}) = 0 | err_{S_1}(h_{S_1}) > \epsilon) \\
&\leq (1-\epsilon)^{n-k},
\end{aligned}$$

   i.e., the probability of event $A_{i_1,\dots,i_k}$ is at most $(1-\epsilon)^{n-k}$. □

   (b) Now use this to prove the guarantee in the displayed equation above.

   Proof. For S with $|S| = n$, the number of subset $S' \subseteq S$ such that $|S'| = k$ is $\binom{n}{k}$. By union bound, we have

$$\begin{aligned}
&\mathbb{P}_{S \sim D^n}\left(\exists S' \subseteq S, |S'| = k, \text{ such that } h_{S'} \text{ has 0 error on } S - S' \text{ but true error } > \epsilon\right) \\
=&\mathbb{P}_{S \sim D^n}\left(\bigcup_{i_1,\dots,i_k} A_{i_1,\dots,i_k}\right) \\
\leq&\binom{n}{k}\mathbb{P}_{S \sim D^n}\left(A_{i_1,\dots,i_k}\right) \\
=&\binom{n}{k}(1-\epsilon)^{n-k} \\
\leq&\binom{n}{k}e^{-(n-k)\epsilon}.
\end{aligned}$$

Solution (cont.)

When n is much larger than k, we have

$$\binom{n}{k} \approx \frac{(n/k - 0.5)^k e^k}{\sqrt{2\pi k}} \simeq O(n^k)$$

from https://en.wikipedia.org/wiki/Binomial_coefficient#n_much_larger_than_k.
Let $n^k e^{-(n-k)\epsilon} \leq \delta$, we have

$$\begin{aligned} k \ln n - (n-k)\epsilon &\leq \ln \delta \\ n\epsilon &\geq k \ln n + k\epsilon + \ln \frac{1}{\delta} \\ n &\geq \frac{1}{\epsilon}\left(k \ln n + k\epsilon + \ln \frac{1}{\delta}\right) \end{aligned}$$

$\square$