TTIC 31260 - Algorithmic Game Theory (Winter 2026)

Homework # 1 **Solutions**

Exercises:

1. **Go with your strengths?** Recall that in the penalty-shot game the shooter can shoot left or shoot right, and the goalie can dive left or dive right. If the goalie guesses correctly then the shot is blocked, else the ball goes in and it is a goal. Now, suppose the shooter is wildly inaccurate on the left, so that even if the goalie dives right, the ball still has only a 25% chance of going in. I.e., the payoff matrix R for the shooter (the row player) looks like:

   Goalie

Shooter	left	right
left	0	1/4
right	1	0

   (The payoff matrix C for the goalie equals $-R$).

   (a) What is the minimax optimal strategy for the shooter, and what is the minimax optimal value?

   Solution: Say p is the probability on left and $1 - p$ is the probability on right. Then the shooter gets payoff $1 - p$ if the goalie dives left and $p/4$ if the goalie dives right. Since one is decreasing with p and one is increasing with p, this means the minimum is maximized when the two are equal. This solves to $p = 4/5$. The minimax optimal value is $1/5$.

   (b) Is there anything perhaps counterintuitive about how weakening the shooter on the left affected its minimax optimal strategy?

   You'd think the shooter might want to go with its strength, but in fact it's the opposite, increasing probability on its weakness.

   (c) What is the minimax optimal strategy for the goalie?

   By similar reasoning, this solves to diving left with probability $1/5$ and diving right with probability $4/5$.

2. **Exchangability.** Suppose that (p, q) and (p', q') are both Nash equilibria for some game given by payoff matrices (R, C).

   (a) Must (p, q') be a Nash equilibrium? If so, give a proof; if not, give a counterexample.

   Solution: no. E.g., consider the walking-on-the-sidewalk game. LL is an equilibrium and RR is an equilibrium, but LR is not.

(b) Suppose the game is zero-sum ($C = -R$). Must (p, q') be a Nash equilibrium? If so, give a proof; if not, give a counterexample.

Solution: yes. Since it is zero-sum, we know p and p' are both minimax optimal strategies for the Row player, guaranteeing expected payoff at least the value v of the game no matter what the Column player plays, and q and q' are both minimax optimal for the Column player, guaranteeing expected payoff at least $-v$ no matter what the Row player plays. Therefore, under (p, q') the row player has expected gain exactly v and neither player can perform better by deviating, so it is a Nash equilibrium.

3. **Game-theoretic RWM (multiplicative weights) analysis.** In class we described the RWM algorithm in the context of predicting from expert advice, and briefly sketched how the analysis extended to the game-theoretic setting where the algorithm is *selecting an action* rather than combining predictions. Here we want you to go ahead and do that analysis.

Specifically, assume you have N action choices (rows in a matrix game). The algorithm maintains a weight w_i for each row i (initialized to $w_i = 1\ \forall i$) and defines probability distribution $p_i = w_i/W$ where $W = \sum_i w_i$. To make this a little different from the discussion in class, let's assume that payoffs are *gains* in the range $[0, 1]$ and that if action i receives a gain of g_i then the algorithm updates w_i using

$$w_i \leftarrow w_i(1 + \epsilon g_i),$$

where $0 < \epsilon \leq 1$ is an input to the algorithm.

Let $G_i = \sum_t g_i^t$ denote the cumulative gain of action i over all time steps t so far (superscript t denotes the time-step), and let $Q_i = \sum_t (g_i^t)^2 \leq G_i$. Let G_{ALG} denote the expected cumulative gain of the above algorithm. Prove that

$$G_{ALG} \geq \max_i \left[G_i - \epsilon Q_i/2 - (\ln N)/\epsilon\right] \geq \max_i \left[G_i(1 - \epsilon/2) - (\ln N)/\epsilon\right].$$

Feel free to look at the proof of the closely-related Theorem 4.6 in the book if you like. Also you will want to use the fact that for $x \in [0, 1]$ we have $x - x^2/2 \leq \ln(1 + x) \leq x$.

Solution: We have:

$$W^{t+1} = W^t + \epsilon \sum_i w_i^t g_i^t = W^t + \epsilon W^t \sum_t p_i^t g_i^t = W^t(1 + \epsilon g_{ALG}^t)$$

where g_{ALG}^t is the expected gain of the algorithm at time t. Chaining together and using the fact that $W^0 = N$ we get $W^T = N \prod_t (1 + \epsilon g_{ALG}^t)$. Taking logs,

$$\ln(W^T) = \ln(N) + \sum_t \ln(1 + \epsilon g_{ALG}^t) \leq \ln(N) + \epsilon G_{ALG}.$$

For the lower bound, for any expert i we have:

$$\ln(W^T) \geq \ln(w_i^T) = \sum_t \ln(1 + \epsilon g_i^t) \geq \sum_t \epsilon g_i^t - \sum_t \epsilon^2 (g_i^t)^2/2 = \epsilon G_i - \epsilon^2 Q_i/2.$$

Putting these together and dividing by ϵ we have:

$$G_{ALG} \geq G_i - \epsilon Q_i/2 - (\ln N)/\epsilon \geq G_i(1-\epsilon/2) - (\ln N)/\epsilon$$

as desired.

Problems:

4. **On approximate Nash equilibria.** Consider a two-player n-by-n general-sum game. Recall that in a Nash equilibrium (p, q), for each i s.t. $p_i > 0$ we have $e_i^T Rq = \max_{i'} e_{i'}^T Rq$ and similarly for each j s.t. $q_j > 0$ we have $p^T Ce_j = \max_{j'} p^T Ce_{j'}$.

   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. Even more stringently, define a pair of distributions (p, q) to be a *well-supported ϵ-Nash equilibrium* if each row i with $p_i > 0$ satisfies $e_i^T Rq \geq \max_{i'} e_{i'}^T Rq - \epsilon$ and each row j with $q_j > 0$ satisfies $p^T Ce_j \geq \max_{j'} p^T Ce_{j'} - \epsilon$.[1]

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

   Hint: you will want somewhere to use Hoeffding's inequality, which says that if $X = \frac{1}{m}\sum_{i=1}^m X_i$ where the X_i are independent $[0, 1]$-valued random variables, then

   $$\Pr[|X - \mathbf{E}[X]| > \epsilon] \leq 2e^{-2m\epsilon^2}.$$

   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.

   Solution: Consider some Nash equilibrium (p, q). Let S be a (multi-)set of k rows selected iid from p, and let T be a multi-set of k rows selected iid from q. Let U_S denote the uniform distribution over S and let U_T denote the uniform distribution over T. The claim is that $k = O(\frac{1}{\epsilon^2}\log n)$ is sufficient so that with high probability, the pair (U_S, U_T) is an ϵ-Nash equilibrium (so such a pair must exist). In particular, by Hoeffding bounds, this value of k is sufficient so that with high probability, for every column c, its average payoff over the rows in S is within $\pm\epsilon/2$ of its expected payoff with respect to the distribution p. Similarly, with high probability, for every row r, its average payoff over the columns in T is within $\pm\epsilon/2$ of its expected payoff with respect to the distribution q. We can set the constant in the big-Oh such that the failure probability of any given row or column is $o(1/n)$ so the overall probability of failure is $o(1)$. So long as both conditions hold, this implies that the pair (U_S, U_T) has the property that each player has at most ϵ incentive to deviate.

[1] So the difference is that in an ϵ-Nash equilibrium, players can put a small amount of probability on very bad actions, whereas in a well-supported ϵ-Nash they cannot.

5. **Tracking a moving target.** Here is a variation on the deterministic Weighted-Majority algorithm, designed to make it more adaptive.

   (a) Each expert begins with weight 1 (as before).

   (b) We predict the result of a weighted-majority vote of the experts (as before).

   (c) If an expert makes a mistake, we penalize it by dividing its weight by 2, but *only* if its weight was at least 1/4 of the average weight of experts.

   Prove that in any contiguous block of trials (e.g., the 51st example through the 77th example), the number of mistakes made by the algorithm is at most $O(m + \log N)$, where m is the number of mistakes made by the best expert *in that block*, and N is the total number of experts.

   Solution: Let W_{init} be the total weight at the beginning of the interval and W_{final} be the total weight at the end of the interval.

   First, notice that all weights are at least $1/8$ of the average. We can see this by induction: the average never increases, so the statement holds for weights that were not lowered in the last round. Also, if a weight was lowered, then it must have been at least $1/4$ of the old average, so it is now at least $1/8$ of the old average which is at least $1/8$ of the new average.

   This means that the weight of the best expert at beginning of the interval is at least $W_{init}/(8n)$, and therefore by end of the interval it is at least $(1/2)^m W_{init}/(8n)$.

   Also, on each mistake, at most $W/4$ of the total weight is fixed. So at least $(W/2 - W/4) = W/4$ gets cut in half. In other words, $W/8$ is removed from the total weight. This means $W_{final} < W_{init}(7/8)^M$.

   The bound results from solving $(1/2)^m W_{init}/(8n) \leq W_{init}(7/8)^M$.