Instructions

This quiz comprises 5 short problems. Attempt all questions and submit your solutions on Canvas on the quizzes section. Please use Canvas, do not turn in hard (nor soft) copies.

This homework is to be completed **individually** and is due in ~ 3 weeks. You are encouraged to complete the assignment without AI usage, though I cannot and will not police or enforce this.

Important erratas are given in red.

- Q2: $||x||_1$ is corrected to read $||x||_0$ in (1).

1 Add in linear combinations of other actions

Let $G = (A, B), A, B \in \mathbb{R}^{n \times m}$ be a general sum game with some Nash equilibrium (x^*, y^*). Let $\mathbf{a}_{r_i} \in \mathbb{R}^{m \times 1}$ be the i-th row of A and $\mathbf{a}_{c_j} \in \mathbb{R}^{n \times 1}$ be the j-th column of A

$$A = \left[\begin{array}{ccc} a_{11} & \cdots & a_{1m} \\ \hline a_{21} & \cdots & a_{2m} \\ \hline \vdots & \ddots & \vdots \\ \hline a_{n1} & \cdots & a_{nm} \end{array}\right] = \left[\begin{array}{c} \mathbf{a}_{r_1}^\top \\ \hline \mathbf{a}_{r_2}^\top \\ \hline \vdots \\ \hline \mathbf{a}_{r_n}^\top \end{array}\right],$$

$$A = \left[\left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{array}\right]\right] = \left[\begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array}\right],$$

Furthermore, we define the payoff vectors

$$\mathbf{a}_{r_{\text{nash}}} = \sum_{i \in [n]} x^*(i) \cdot \mathbf{a}_{r_i}, \quad \mathbf{a}_{c_{\text{nash}}} = \sum_{j \in [m]} y^*(j) \cdot \mathbf{a}_{c_j}$$

as the mixture of rows (resp. columns) weighted according to x^* (resp y^*). Define $\mathbf{b}_{r_i}$, $\mathbf{b}_{c_j}$, $\mathbf{b}_{r_{\text{nash}}}$ and $\mathbf{b}_{c_{\text{nash}}}$ similarly, but using B instead.

We add a single action to each player, which is to play that player's component of the NE, x^* or y^* respectively. This yields a new game $G = (A', B'), A', B' \in \mathbb{R}^{(n+1) \times (m+1)}$. The new action for P1 (resp. P2) is the linear combination of all rows (resp. columns) weighted according to x^* (resp. y^*). Specifically, the new game's payoff matrixes are

$$A' = \begin{bmatrix} A & \mathbf{a}_{c_{\text{nash}}} \\ \mathbf{a}_{r_{\text{nash}}}^\top & {x^*}^\top A y^* \end{bmatrix}, \quad B' = \begin{bmatrix} B & \mathbf{b}_{c_{\text{nash}}} \\ \mathbf{b}_{r_{\text{nash}}}^\top & {x^*}^\top B y^* \end{bmatrix},$$

In this new game G', is the pure strategy profile of both players playing the new action a Nash equilibrium? Select *one* option.

(A) Always Nash for all zero-sum games, but not necessarily for general-sum games.

(B) Always Nash, including for general-sum games.

(C) Not necessarily Nash, even if game is zero-sum.

(D) None of the above hold.

Solution. B. Simply use the definition of Nash equilibrium to see that deviating to one of original n (resp. m) actions will not yield a strictly better expected payoff for P1 (resp. P2).

2 Sparse strategies and equilibria

Consider the *zero-sum game* with payoff matrix $A \in \mathbb{R}^{n\times m}$. Following our convention, you (the row player, P1) is the minimizing player. Denote the value of this game by $v \in \mathbb{R}$.

$$v = \min_{x\in\Delta_n} \max_{y\in\Delta_m} x^\top Ay.$$

Consider the following modification where you have the constraint that you are only able to commit to k-sparse strategies, i.e., you may only randomize among $1 \le k \le n$ actions and announce this random choice publicly. Mathematically, we are concerned with the following problem

$$v' = \min_{\substack{x\in\Delta_n \\ ||x||_0 \le k}} \max_{y\in\Delta_m} x^\top Ay. \tag{1}$$

Let the optimal value of this modified game be $v' \in \mathbb{R}$. Which of the following statements are true? Select all which apply.

(A) For all $1 \le k \le n$, $v' \ge v$, i.e., the modified game has is always no better for the minimizing than before.

(B) If $v' = v$ when $k = 1$, then the original (unrestricted) game is guaranteed to have a pure strategy Nash equilibrium.

(C) If $v' \ne v$ when $k = 1$ then the original (unrestricted) game is guaranteed to have no pure strategy Nash equilibrium.

(D) Suppose $n > m$, i.e., A is "tall and skinny". Then if $k = m$ we are guaranteed that $v = v'$. For simplicity, you may assume that the game is nondegenerate, i.e., every there are at most z many pure best responses of P1 to any mixture over z actions of P2, and vice versa.

Note: Technically speaking, the solution to (1) is not an equilibrium in the traditional sense.

Solution.

- A is true. If you treat the inner maximization problem as a function of x, we can see that P1 has a strictly stronger constraint and hence it's optimal cannot be better than before.
- B is False. Consider
$$A = \begin{bmatrix} 5 & 5 \\ 10 & 0 \\ 0 & 10 \end{bmatrix}.$$
Clearly, solution to (1) has $x = [1, 0, 0]$. This is also (one of) the x^*, a Nash to the original game. However, the *only* y^* is $[0.5, 0.5]$. There is no pure equilibrium.
- C is True. We can prove the contrapositive. If there was a pure NE, then $v' = v$ clearly, since the sparse commitment to x for the modified game is just to play that action.
- D is True. For those who are familiar with linear programs, this is straightforward from the linear programming formulation. If not, we can use the nondegeneracy assumption. Take any Nash of the original game. That NE for y^* has *at most* a support of $k = m < n$, hence there are at most m best responses of $P1$. Thus, by definition of Nash, x^* also has a support of at most size k and that gives us to the optimal x in the modified game.

3 Making an unfair deal

Consider the following zero-sum game with payoff matrix A. Following our convention, you (the row player, P1) is the minimizing player. Suppose the value of this game is v. In addition, we suppose the game is win-loss, i.e., $A \in \{-1, 1\}^{n \times m}$, i.e., either you win ($-1$) or lose (1).

You realize that $v > 0$ strictly, that is, even playing the Nash equilibrium you are losing more on average. This seems unfair, so you complain to the game-master. Slightly bemused, the game-master offers you the following deal. You can choose to play as P2 instead of P1 (and become the maximizing player). However, this comes with the caveat that instead of playing simultaneously, you must reveal exactly which action (not randomized) you have chosen before your opponent selects theirs.

Which of the following are true? Select all which apply.

(A) If $v = 1$, then you should always take up the deal.

(B) Suppose that P2 has an action that always wins regardless of what P1 does, i.e., there is a column containing all ones. Then you should take up the deal.

(C) Suppose it is desirable (i.e., you do *strictly better* than before as P1) that you take up the deal. Then in the original game, P2 must have an action that always wins no matter what action P1 takes.

(D) Suppose it is desirable that you take up the deal. Then in the original game, P1 must have an action that always loses (for himself) no matter what action P2 takes.

Solution.

- A is true. If v is the value of the game, then there exists some y^* such that Ay^* is the all-ones vector. Because if it was not, then it would not solve the max-min problem.

  But, Ay^* is a convex combination of some columns, whose maximum value is 1. Thus, for it to be the all one's vector, all columns that are included with positive proportion must be all ones; if there was a column with one entry strictly less than one, then the convex combination would have the index corresponding to that entry be strictly less than one.

  This means that there is an all ones column, and if you switch positions to be P2, you are guaranteed a win by playing that column all the time.

- B is true, clearly, since 1 is the highest payoff obtainable. Again, we just play that column all the time.

- C is true. If it were desirable, then you as P2 must *at the very least* be able to guarantee a strictly higher payoff than -1 (which is the lowest possible payoff in the original game). But since the opponent chooses after you choose an action and the game is win-loss, that means playing as P2 must give you exactly 1. Which in turn means that a column in A must be all ones, which is a pure strategy that wins no matter what P1 plays.

- D is false. Consider

$$A = \begin{bmatrix} -1 & 1 \end{bmatrix}.$$

Of course, the value of this game is 1, but if P2 was silly and played the first action, P1 would still win.

4 Online learning and self-play

In class, we said that if 2 players engaged in self-play and each played a strategy that enjoyed sublinear regret, then the average strategy converges to a Nash equilibrium (for zero-sum games). We also covered a few regret minimizers.

Let us suppose P1 played a regret minimizer (e.g., Hedge, RM, RM+), giving iterates $x^{(t)}$ but P2 does not. Instead, P2 simply best responds by playing $y^{(t)} \in \mathrm{BR}_2(x^{(t)})$, i.e., it is allowed to observe $x^{(t)}$ before best responding. Tiebreaks for the best response can be chosen arbitrarily. Which of the following is true? Pick all which apply.

(A) The regret of the first learner is always non-positive for *all* iterations $t > 1$.

(B) The regret of the second learner is always non-positive for *all* iterations $t > 1$.

(C) The empirical average of the strategies played by P1 during self-play converges to a (mixed strategy) Nash equilibrium for P1.

(D) The empirical average of the strategies played by P2 during self-play converges to a (mixed strategy) Nash equilibrium for P2.

Note: you may assume we are talking about external regret, which was the notion of regret that we focused on in class so far. We also mean cumulative regret up to time t, not just from the action that is taken at t.

Solution.

- A is False. This is in general not true. One example is the trivial case where $A = [-1, 1]^\mathsf{T}$, so P2 has no choice of action. At timestep 1, both actions are played equally likely, as opposed to the optimal action. So that means one mistake ensures positive cumulative regret forever.
- B is True. Since we are always best responding to individual strategies at each timestep, we are guaranteed zero or even negative regret.
- C is True. Yes. Because the regrets are still sublinear in T for both learners, we can use the same theorem in class to show this is also an approximate Nash.
- D is True. Same as before.

5 Game Solvers

Which of the following statements about game solving/solvers are true? Pick all which apply.

(A) Finding a Nash equilibrium can be done for 2-player general-sum matrix games using mixed-integer linear programming.

(B) The Lemke Howson algorithm may not converge for 2-player general-sum matrix (you may assume symmetric, non-degenerate) games.

(C) Suppose we are using regret matching for both players with simultaneous updates to solve zero-sum games. Except for the first iteration, the iterates $x^{(t)}, y^{(t)}$ for *all* $t > 1$ are such that strictly dominated actions are played with zero probability.

(D) Suppose we are using Hedge/Multiplicative weights for both players with simultaneous updates to solve zero-sum games. You can assume a suitable learning rate is chosen. Except for the first iteration, the iterates $x^{(t)}, y^{(t)}$ for *all* $t > 1$ are such that strictly dominated actions are played with zero probability.

Solution.

- A is true. We covered this in class.
- B is false, we covered this in class.
- C is false. Suppose $A = [100, 0, \epsilon]^\mathsf{T}$, where $\epsilon > 0$ and is small. Clearly the best strategy is for P1 to always play action 2. Note P2 has only one trivial action. In the first iteration, every action is played with probability 1/3. So, the regret from not playing action 2 and 3 is both highly positive, and in iteration 2 action 3 will be played with almost probability 0.5, even though it is dominated by action 2.
- D is false. Hedge/multiplicative weights always have every action played with strictly non-zero probability.