

ECE 586BH: Problem Set 3: Problems and Solutions

Fictitious play dynamics, minimum regret learning from forecasters

Due: Tuesday, October 16, at beginning of class**Reading:** Course notes, Chapter 3

1. [On potential games with two players]

- (a) Consider a two player normal form game such that for some $(0, 0) \in S \triangleq S_1 \times S_2$, the following holds for all $(a, b) \in S$:

$$u_1(a, 0) - u_1(0, 0) + u_2(a, b) - u_2(a, 0) = u_2(0, b) - u_2(0, 0) + u_1(a, b) - u_1(0, b).$$

Does the game necessarily have a potential function?

Solution: If it does have a potential function Φ then we could set $\Phi(0, 0) = 0$ without loss of generality, and also it would have to satisfy the following, which completely specifies the function:

$$\Phi(a, b) = u_1(a, 0) - u_1(0, 0) + u_2(a, b) - u_2(a, 0).$$

By the given assumption, it follows that $\Phi(a, b) = u_2(0, b) - u_2(0, 0) + u_1(a, b) - u_1(0, b)$. Let's try to prove Φ is a potential function. Let (c, d) and (c', d) be any two strategy profiles differing only in the strategy of player 2. Then

$$\Phi(c, d) = u_1(c, 0) - u_1(0, 0) + u_2(c, d) - u_2(c, 0) \text{ and}$$

$$\Phi(c, d') = u_1(c, 0) - u_1(0, 0) + u_2(c, d') - u_2(c, 0) \text{ so}$$

$\Phi(c, d) - \Phi(c, d') = u_2(c, d) - u_2(c, d')$ as required. Similarly, if (c, d) and (c', d) are any two strategy profiles differing only in the strategy of player 1, then by the alternative expression for Φ ,

$$\Phi(c, d) = u_2(0, d) - u_2(0, 0) + u_1(c, d) - u_1(0, d) \text{ and}$$

$$\Phi(c', d) = u_2(0, d) - u_2(0, 0) + u_1(c', d) - u_1(0, d) \text{ so that}$$

$\Phi(c, d) - \Phi(c', d) = u_1(c, d) - u_1(c', d)$ as required. Hence, Φ is a potential and the game is a potential game.

- (b) Consider a two-person zero sum game such that $u_2(a, b) = \ell(a, b)$ for all $(a, b) \in S$. Thus, $u_1(a, b) = -\ell(a, b)$. Express in terms of ℓ what it means for the game to be a potential game. Simplify as much as possible.

Solution: Part (a) identifies a necessary and sufficient condition on u_1 and u_2 for a two person game to be a potential function (at least if there is a common action 0, which we can achieve by relabeling actions). Expressing that condition in terms of ℓ gives

$$-\ell(a, 0) + \ell(0, 0) + \ell(a, b) - \ell(a, 0) = \ell(0, b) - \ell(0, 0) - \ell(a, b) + \ell(0, b).$$

or equivalently:

$$\ell(a, b) = f(a) + g(b)$$

where $f(a) = \ell(a, 0) - \ell(0, 0)$ and $g(b) = \ell(0, b)$ which means the game decouples into separate optimization problems for the two players.

2. [Best response dynamics for a three person game]

Consider the normal form game with $I = \{1, 2, 3\}$, $S_i = \{0, 1\}$ for each i , and $u_i(s_1, s_2, s_3) = 1$ if the action s_i of player i is different from the actions of both of the other players, and $u_i(s_1, s_2, s_3) = 0$ otherwise. For brevity, we specify a mixed strategy for a single player i by a number $p_i \in [0, 1]$, that indicates the probability the player selects action 1.

- (a) Find the best response set $B(p, q)$ for a player, which is the set of mixed strategies that maximize the (expected) payoff of the player, given the other two players are independently using the mixed strategies p and q , respectively.

Solution: The payoff of a player, assuming it uses strategy $x \in [0, 1]$, against other players using p and q respectively, is $x(1-p)(1-q) + (1-x)pq = x(1-p-q) + pq$ which is maximized over $x \in [0, 1]$ if $x \in B(p, q)$, where

$$B(p, q) = \begin{cases} \{0\} & \text{if } p + q > 1 \\ \{1\} & \text{if } p + q < 1 \\ [0, 1] & \text{if } p + q = 1. \end{cases}$$

- (b) Identify all Nash equilibria in mixed strategies, keeping in mind that mixed strategies include pure strategies as special cases.

Solution: $(0, 1, x)$ is a Nash equilibrium for any $x \in [0, 1]$, and any permutation of such vector is also a Nash equilibrium. In addition $(0.5, 0.5, 0, 5)$ is a Nash equilibrium.

- (c) Identify all trembling hand perfect equilibria.

Solution: Any NE consisting of fully mixed strategies is THP, so $(0.5, 0.5, 0, 5)$ is THP. A strategy of the form $(0, 1, x)$ with $x \notin \{0, 1\}$ is THP, because the sequence of fully mixed strategies $(1/n, 1 - 1/n, x)$ with n so large that $1/n \leq \min\{x, 1 - x\}$, is such that the strategies in $(0, 1, x)$ are best responses to pairs of strategies from $(1/n, 1 - 1/n, x)$. In other words, $0 \in B(1 - 1/n, x)$, $1 \in B(1/n, x)$, and $x \in B(1/n, 1 - 1/n)$. Similarly, $(0, 1, 0)$ is THP (consider the sequence $(1/n, 1 - 1/n, 1/n)$ for $n \geq 2$) and $(0, 1, 1)$ is THP. Thus, all the NE found in part (b) are THP.

- (d) Suppose the players use synchronous best response dynamics. To be definite, assume that if either action has the same payoff for a player against the strategies of the other two players in the previous round, the player selects by flip of a fair coin. Determine the long run limiting payoff per round for each player.

Solution: Suppose the vector of pure actions used in some round is $s = (s_1, s_2, s_3) \in \{0, 1\}^3$. If s is an all equal vector, i.e. either $(0, 0, 0)$ or $(1, 1, 1)$, then the state will alternate between the two all equal vectors forever after. If s is not an all equal vector, then precisely one player wins in the round, so for the next round two players will select a pure strategy by flipping a fair coin. By the symmetry between 0 and 1, the state in the next round will be an all equal vector with probability $1/4$. Thus, if, the initial payoff vector is an all equal vector, then the payoffs for all three players will be zero for all rounds. Otherwise, the number of rounds until an all equal vector is played has the geometric distribution with parameter $1/4$, which is finite with probability one. Thus, the long run limiting payoff per round for each player is zero with probability one.

- (e) Is this a potential game? Justify your answer.

Solution: Yes. Let's seek a potential function $\Phi : \{0, 1\}^3 \rightarrow \mathbb{R}$. Without loss of generality we set $\Phi(0, 0, 0) = 0$. Since $u_1(1, 0, 0) - u_1(0, 0, 0) = 1$ it must be that $\Phi(1, 0, 0) = 1$. Similarly, $\Phi(0, 1, 0) = \Phi(0, 0, 1) = 1$. Since $u_1(1, 0, 0) - u_1(1, 1, 0) = 0$ it must also be that

$\Phi(0, 1, 1) = \Phi(1, 0, 1) = \Phi(1, 1, 0) = 1$. Finally, it must be that $\Phi(1, 1, 1) = 0$. Checking all single changes to strategy profiles, we find that $\Phi(s) = 0$ if $s_1 = s_2 = s_3$ and $\Phi(s) = 1$ otherwise, is a potential. (It implies that asynchronous strict improvement finds a NE after at most one change in strategy profile.)

3. [Regularized fictitious play for a three player game]

For the game of the previous problem, find and simplify as much as possible the equations for regularized fictitious play, including the ordinary differential equation, by modifying the equations from Section 3.4 of the notes. (Suppose for each player that the other two players are selecting actions independently according to their current state distributions.) The differential equation should involve the parameter $\tau > 0$. Numerically identify the equilibrium points and identify which ones are asymptotically stable, for $\tau = 1$ and for $\tau = 0.1$. A bit of Python code that could save you some time is available in a Python notebook file available at <http://courses.engr.illinois.edu/ece586GT/fa2018/pointer.html>.

Solution: The regularized and time-scaled ode model for continuous time fictitious play for this example can be written as a 3-dimensional ode:

$$\dot{q}_i(t) = \beta(q_{-i}(t)) - q_i(t) \text{ for } i \in \{1, 2, 3\}$$

where $\beta : [0, 1]^2 \rightarrow [0, 1]$ is defined so that $\beta(q_2, q_3)$ is the soft arg max of the vector of payoffs for player 1 given by $((1 - q_2)(1 - q_3), q_2 q_3)$. In other words (remembering that we are denoting mixed strategies for $S_i = \{0, 1\}$ by specifying the probability of using action 1),

$$\begin{aligned} \beta(q_2, q_3) &= \frac{e^{(1-q_2)(1-q_3)/\tau}}{e^{(1-q_2)(1-q_3)/\tau} + e^{q_2 q_3/\tau}} \\ &= \frac{e^{(1-q_2-q_3)/\tau}}{e^{(1-q_2-q_3)/\tau} + 1} \end{aligned}$$

By using the change of variable $q_i = \frac{1}{2} + x_i$ for $x_i \in [-0.5, 0.5]$ the first differential equation becomes

$$\dot{x}_1 = \frac{1}{2} \tanh\left(-\frac{x_2 + x_3}{2\tau}\right) - x_1,$$

and the equations for $\dot{x}_2$ and $\dot{x}_3$ are similar.

We explored the dynamics numerically for $\tau = 0.1$ and $\tau = 1.0$. For $\tau = 1.0$ we found the equilibrium $(0, 0, 0)$ to be globally stable. Apparently for such a large value of τ , the regularization favors high entropy to lead to such stability.

For $\tau = 0.1$ (and smaller τ) the players like to move apart from each other. We found that up to permutation of the three players, the only stable equilibrium has the form $(a, 0, -a)$ where

$$a = \frac{1}{2} \tanh\left(\frac{a}{2\tau}\right)$$

If the x 's are initially unequal their trajectories never cross. In addition, $(0, 0, 0)$ is an unstable equilibrium and is reached if initially all three coordinates of x are equal. If initially the second and third coordinates of x are equal and smaller than the first coordinate, the trajectory

converges to the vector of the form $(x_1, x_2, x_3) = (b, -c, -c)$ with $0 < c < b < 0.5$ determined by

$$b = \frac{1}{2} \tanh\left(\frac{2c}{2\tau}\right)$$

$$c = \frac{1}{2} \tanh\left(\frac{b-c}{2\tau}\right)$$

Such vector is unstable – a small perturbation that makes x_2 larger than x_3 leads to convergence to $(a, 0, -a)$.

4. **[Hannan consistent estimators for a three player game]**

Consider further the game of the previous two problems.

- (a) Suppose player 1 uses a Hannan consistent estimator, where the two experts correspond to actions 0 and 1 for the player. What is the minimum guaranteed long-term payoff of the player, assuming arbitrary behavior of the other two players. **Solution:** Zero. If player 2 keeps playing 0 and player 3 keeps playing 1 then player 1 can never get a positive payoff.
- (b) Suppose all three players use Hannan consistent estimators. Show that the long term sum of payoffs of players 1 through 3 per time slot is greater than or equal to $3/4$. (Hint: There are two experts per player and Hannan consistency means each player's average payoff per unit time is greater than or equal to the payoffs of both of his/her experts.)

Solution: Let η_1, η_2, η_3 denote the payoffs per unit time for the six players. Let $1 - x$ denote the fraction of time no player wins and x denote the fraction of time that one player wins. Then $\eta_1 + \eta_2 + \eta_3 \geq x$. Each player has two corresponding experts so there are six experts. If one player wins then one of the six experts also wins, namely, the expert of the winning player that played the same action as the player. A fraction $1 - x$ of the time, no player wins, because all three players play the same action. In those time slots, three of the experts win, namely, one for each player that did not agree with the player. So if $w_{i,j}$ denotes the win rate of expert j for player i , then the sum of the six w 's is $x + 3(1 - x) = 3 - 2x$. By Hannan consistency of the estimators, $x_i \geq \max\{w_{i,0}, w_{i,1}\}$ for each i , so that

$$\eta_1 + \eta_2 + \eta_3 \geq \sum_{i=1}^3 \max\{w_{i,0}, w_{i,1}\} \geq \frac{1}{2} \sum_{i=1}^3 w_{i,0} + w_{i,1} = \frac{3 - 2x}{2}.$$

We thus have $\eta_1 + \eta_2 + \eta_3 \geq \max\left\{x, \frac{3-2x}{2}\right\} \geq 3/4$, where the second inequality holds with equality for $x = 3/4$.

5. **[Strong law of large numbers for discrete-time martingales with bounded increments]**

Show that if $(D_t : t \geq 0)$ is a martingale difference sequence such that $|D_t| \leq 1$ for all t , then the corresponding martingale defined by $X_n = D_1 + \dots + D_n$ satisfies $\limsup_{n \rightarrow \infty} \frac{X_n}{2\sqrt{n \ln n}} \leq 1$ almost surely (i.e. with probability one). (Hint: Use the Azuma-Hoeffding inequality and the Borel Cantelli lemma – both in the notes.) In particular, $\frac{X_n}{n} \rightarrow 0$ almost surely. (This is used to help establish the existence of Hannan consistent estimators for players of a finite game. See the paragraph following Corollary 3.26.

Solution: By the Azuma-Hoeffding inequality with $A_t = -1$ and $B_t = 1$, for $n \geq 1$,

$$\mathbb{P}\{X_n \geq \gamma\} \leq e^{-\frac{\gamma^2}{2n}}$$

The problem statement suggests we take $\gamma = 2\sqrt{n \ln n}$, which yields

$$\mathbb{P}\{X_n \geq 2\sqrt{n \ln n}\} \leq e^{-2 \ln n} = \frac{1}{n^2}$$

Since $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$, the Borel-Cantelli lemma implies $\mathbb{P}\left\{\{X_n \geq 2\sqrt{n \ln n}\} \text{ i.o.}\right\} = 0$, so that $\limsup_{n \rightarrow \infty} \frac{X_n}{2\sqrt{n \ln n}} \leq 1$ a.s.

6. **[Blackwell approachability for a simple game with vector valued payoff]**
Consider a two player game with strategy spaces

$$S_1 = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} \text{ and } S_2 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\},$$

such that if player 1 selects $x \in S_1$ and player 2 selects $y \in S_2$, the payoff vector for player 1 is $x + y$.

- (a) An arbitrary half space H in the plane can be expressed as $H = \{v : v \cdot u \leq c\}$, where u is a unit length vector in $\mathbb{R}^2$ and $c \in \mathbb{R}$. Under what condition on u and c is this half space approachable (for player 1)?

Solution: The half space H is approachable if and only if the set $(-\infty, c]$ is approachable for the auxiliary game with payoff $x \cdot u + y \cdot u$, where player 1 is the minimizer and player 2 is the maximizer. Player 1 can make $x \cdot u$ assume any one of the four values $u_1, -u_2, u_2, -u_1$ and player 2 can make $y \cdot u$ assume any value of the form $\pm(u_1 + u_2)$. Player 1 has a dominant strategy x resulting in $x \cdot u = -\max\{|u_1|, |u_2|\}$. Player 2 has a dominant strategy y resulting in $y \cdot u = |u_1 + u_2|$. The value of the auxiliary game is thus $|u_1 + u_2| - \max\{|u_1|, |u_2|\}$, and H is approachable if and only if $c \geq |u_1 + u_2| - \max\{|u_1|, |u_2|\}$. (An equivalent condition is $c \geq \text{sign}(u_1 u_2) \min\{|u_1|, |u_2|\}$.)

- (b) Let $B(r)$ denote the disk in $\mathbb{R}^2$ of radius r . Find the minimum value of r so that $B(r)$ is approachable (for player 1).

Solution: By Blackwell's theorem, $B(r)$ is approachable if and only if any halfspace containing $B(r)$ is approachable. A halfspace $H = \{x : x \cdot u \leq c\}$ with u being a unit length vector contains $B(r)$ if and only if $c \geq r$. Thus, the minimum r is equal to the minimum value of c in part (a) such that H is approachable for any unit length vector u . So $r = \max\{|u_1| + |u_2| - \max\{|u_1|, |u_2|\}\}$, where the maximum is over unit length vectors u . By changing the sign of u_2 if necessary, we can assume without loss of generality that u_1 and u_2 have the same sign, in which case $|u_1 + u_2| - \max\{|u_1|, |u_2|\} = \min\{|u_1|, |u_2|\}$. This minimum is maximized by selecting $|u_1| = |u_2| = \frac{1}{\sqrt{2}}$. So the minimum value of r is $\frac{1}{\sqrt{2}}$.