University of Illinois **Fall 2017**

ECE 586BH: Problem Set 3: Problems and Solutions
Fictitious play dynamics, minimum regret learning from forecasters

Due: Tuesday, October 17, at beginning of class
Reading: Course notes, Chapter 3

1. **[Which symmetric games are potential games?]**
A two player normal form game $G = (\{1,2\}, S_1, S_2, u_1, u_2)$ is symmetric if $S_1 = S_2$ and $u_1(i,j) = u_2(j,i)$ for any $i, j \in S_1$. The following definition makes sense because u_2 is determined by u_1 for such a game. Define a function u_1 on a set $S_1 \times S_1$ to have property P if it is the payoff function for player 1 in a symmetric two person potential game.

   (a) If there is a function g such that $u_1(i,j) = g(i)$ for all $i, j \in S_1$, does u_1 have property P ? If so, give the potential function. If not, give a counter example.
   Solution: Property P means there exists a potential function such that $u_1(i',j) - u_1(i,j) = \Phi(i',j) - \Phi(i,j)$ and $u_2(j,i') - u_2(j,i) = \Phi(j,i') - \Phi(j,i)$. By the symmetry of the game, the latter condition is the same as $u_1(i',j) - u_1(i,j) = \Phi(j,i') - \Phi(j,i)$. So to summarize, property P is equivalent to the existence of a function Φ such that:

   $$u_1(i',j) - u_1(i,j) = \Phi(i',j) - \Phi(i,j) = \Phi(j,i') - \Phi(j,i)$$

   for all $i, i', j \in S_1$.
   If $u_1(i,j) = g(i)$ the condition becomes $g(i') - g(i) = \Phi(i',j) - \Phi(i,j) = \Phi(j,i') - \Phi(j,i)$, which is satisfied for the function $\Phi(i,j) = g(i) + g(j)$. Answer is yes.

   (b) If $u_1(i,j) = H(i,j)$ such that H is symmetric (i.e. $H(i,j) = H(j,i)$ for all $i, j \in S_1$) does u_1 have property P? If so, give the potential function. If not, give a counter example.
   Solution: For the given form of u_1, the condition found in part (a) becomes $H(i',j) - H(i,j) = \Phi(i',j) - \Phi(i,j) = \Phi(j,i') - \Phi(j,i)$, which is satisfied by $\Phi(i,j) = H(i,j)$. Equivalently, u_1 itself is a potential function. Answer is yes.

   (c) For S_1 fixed and a function $u_1 : S_1 \times S_1 \to \mathbb{R}$, does property P hold if and only if there exist $g : S_1 \to \mathbb{R}$ and $H : S_1 \times S_1 \to \mathbb{R}$ such that H is symmetric and $u_1(i,j) = g(i) + H(i,j)$ for all $i, j \in S_1$? Justify your answer. Hint: How would you construct and verify a potential function for the symmetric game, given u_1?
   Solution: Yes. (if) The sum of two functions with property P also has property P, so if u_1 has the form $u_1(i,j) = g(i) + H(i,j)$ for a symmetric H then u_1 has property P by parts (a) and (b). Moreover, a potential function is $\Phi(i,j) = g(i) + g(j) + H(i,j)$.
   (only if) Fix $k \in S_1$. If a potential Φ for the symmetric game existed, it would have to satisfy both the following, obtained by considering the sequence of profiles $(k,k) \to (i,k) \to (i,j)$ and $(k,k) \to (k,j) \to (i,j)$, respectively:

   $$\Phi(i,j) - \Phi(k,k) = (u_1(i,k) - u_1(k,k)) + (u_2(i,j) - u_2(i,k)) = u_1(i,k) - u_1(k,k) + u_1(j,i) - u_1(k,i)$$
   $$\Phi(i,j) - \Phi(k,k) = (u_2(k,j) - u_2(k,k)) + (u_1(i,j) - u_1(k,j)) = u_1(j,k) - u_1(k,k) + u_1(i,j) - u_1(k,j).$$

   The existence of Φ would thus require the right hand sides of the above to be equal, which after some rearrangement yields:

   $$u_1(i,j) + u_1(k,i) - u_1(i,k) = u_1(j,i) + u_1(k,j) - u_1(j,k). \tag{1}$$

Equivalently, if we define $H(i,j) = u_1(i,j) + u_1(k,i) - u_1(i,k)$, then H is a symmetric function. Letting $g(i) = u_1(i,k) - u_1(k,i)$ we see $u_1(i,j) = g(i) + H(i,j)$.

2. **[Doubling trick]**
Proposition 3.18 of the notes on prediction with expert advice provides the upper bound $\max_i R_{i,n} \leq \sqrt{2n \ln N}$ for the exponentially weighted forecaster with parameter $\eta = \sqrt{\frac{2 \ln N}{n}}$. In some situations there is not a single time horizon n known in advance, so it would be nice to have a forecaster and similar maximum regret bound that holds uniformly over time. That can be accomplished by letting η get smaller with time. A simple approach to this problem is to partition time into periods $I_0 = \{1\}$, $I_1 = \{2,3\}$, $\ldots$, $I_k = \{2^k, \ldots, 2^{k+1} - 1\}, \ldots$ doubling in length. Instead of carrying over information from one interval to the next, suppose we simply apply the algorithm used Proposition 3.18 separately for each interval.

   (a) Verify that the proof of Proposition 3.18 shows that a bit more is true, namely, $\max_i R_{i,n'} \leq \sqrt{2n \ln N}$ for any n' with $1 \leq n' \leq n$.
   Solution: Very little modification of the proof is needed. The line just before the statement of the proposition can be changed to
   $\max_i R_{i,n'} \leq \Phi(R'_n) \leq \frac{\ln N}{\eta} + \frac{n'\eta}{2} \leq \frac{\ln N}{\eta} + \frac{n\eta}{2} = \sqrt{2n \ln N}$ for $1 \leq n' \leq n$, where $\eta = \sqrt{\frac{2 \ln N}{n}}$.

   (b) For the forecaster based on the doubling trick, describe the choice of η in each interval and find the smallest constant c so that $\max_i R_{i,n} \leq c\sqrt{2n \ln N}$ for all $n \geq 1$. (Hint: The cumulative regret at $t \in I_k$ is the sum for intervals I_0 through I_k.
   Solution: The length of the kth interval is 2^K for $k \geq 0$, so use the algorithm for fixed n equal to 2^k for the kth interval. So use $\eta = \sqrt{2^{1-k} \ln N}$ for $t \in I_k$. Then, for any $t \in I_k$, $\max_i R_{i,t}$ is less than or equal to the sum of the regret bounds for intervals of length $1, 2, 4, \ldots, 2^k$, or

$$\sqrt{2 \ln N}(2^{0/2} + 2^{1/2} + \ldots + 2^{k/2}) = \sqrt{2 \ln N} \left(\frac{2^{\frac{k+1}{2}-1}}{\sqrt{2}-1} \right).$$

   Thus we need c so that

$$\frac{2^{\frac{k+1}{2}} - 1}{\sqrt{2}-1} \leq c\sqrt{t}$$

   for any $t \in I_k$ for any $k \geq 0$. This inequality will be true for all $t \in I_k$ if it is true for the smallest $t \in I_k$, namely, for $t = 2^k$. Thus, we need c so that

$$\frac{2^{\frac{k+1}{2}} - 1}{\sqrt{2}-1} \leq c2^{\frac{k}{2}}$$

   for $k \geq 0$. For large k, we need $c = \frac{\sqrt{2}}{\sqrt{2}-1}$, or larger. In fact, such c works for all k as we see by dropping the "-1" in the numerator. Multiply above and below by $\sqrt{2}+1$ to get $c = 2 + \sqrt{2} \approx 3.4142$.

3. **[Existence of Hannan consistent estimators]**

(a) 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 $X_n = X_1 + \cdots + X_n$ satisfies $\limsup_{n\to\infty} \frac{X_n}{2\sqrt{n \ln n}} \leq 1$ almost surely (i.e. with probability one). (Hint: Use the Azuma-Hoeffding inequality and the Borel Cantelli lemma – both in the notes.)

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

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

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

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

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 $\limsup_{n\to\infty} \frac{X_n}{2\sqrt{n \ln n}} \leq 1$ a.s.

(b) Combine part (a) and previous problem to show a Hannan consistent estimator exists.

Solution: See the paragraph following Corollary 3.26 in the course notes. That paragraph mentions using Propostion 3.22 for bounding the sum of expected regrets for all n, but we could also use the doubling trick as mentioned in the previous problem for that part.

4. **[One Hannan consistent player]**
Consider a finite two-player zero-sum game such that $\ell : S_1 \times S_2 \to \mathbb{R}$ serves as the loss function of player 1 and reward function of player 2, as usual. Suppose the game is played repeatedly to produce a sequence of outcome tuples $(I_t, J_t)_{t\geq 1}$.

(a) State what it means in this instance for player 1 to use a Hannan consistent strategy.

Solution: By definition, (I_t) is produced by a Hannan connsistent strategy if

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \ell(I_t, J_t) - \min_{i\in S_1} \sum_{t=1}^{n} \ell(i, J_t) \leq 0$$

with probability one, no matter what randomized, causal strategy is used by player 2.

(b) If player 1 uses a Hannan consistent strategy, and if p^* is a limit point of the empirical distribution of plays of player one, $(\hat{p}_n^I)_{n\geq 1}$, then must p^* be a minmax optimal strategy for player 1 for a single play of the game? If yes, give a proof. If no, give a counter example.

Solution: No. For example, suppose $S_1 = S_2 = \{1, 2\}$ and ℓ in matrix form is $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$. Suppose during the sequence of games that player 2 selects $J_t = 1$ for all t. Since player 1 is Hannan consistent, it must be that $\hat{p}_n^I \overset{n\to\infty}{\longrightarrow} (1, 0)$. But $(0, 1)$ is the unique minmax strategy for player 1. (The idea is that player 2 plays poorly during the sequence of games so that player 1 is not forced to settle on a minmax optimal empirical distribution of plays.)

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

$$S_1 = \left\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\-1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\end{pmatrix}\right\} \text{ and } S_2 = \left\{\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}, \begin{pmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{pmatrix}, \begin{pmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}, \begin{pmatrix}-\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{pmatrix}\right\},$$

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 $u_1, -u_2, u_2, -u_2$ and player 2 can make $y \cdot u$ assume any value of the form $\frac{\pm u_1 \pm u_2}{\sqrt{2}}$. 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 $y \cdot u = \frac{|u_1|+|u_2|}{\sqrt{2}}$. The value of the auxiliary game is thus $\frac{|u_1|+|u_2|}{\sqrt{2}} - \max\{|u_1|, |u_2|\}$, and H is approachable if and only if $c \geq \frac{|u_1|+|u_2|}{\sqrt{2}} - \max\{|u_1|, |u_2|\}$.

(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.

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 $r = \max\left\{\frac{|u_1|+|u_2|}{\sqrt{2}} - \max\{|u_1|, |u_2|\}\right\}$, where the maximum is over unit length vectors u. By symmetry, we can assume $u_1 = a$ and $u_2 = \sqrt{1-a^2}$, where $\frac{1}{\sqrt{2}} \leq a \leq 1$, so we find the maximum of $\frac{a+\sqrt{1-a^2}}{\sqrt{2}} - a$ over that range. The derivative is negative throughout the range, so that $a = \frac{1}{\sqrt{2}}$ achieves the maximum value $r = 1 - \frac{1}{\sqrt{2}}$.

6. **[Pursuit evasion game]**
Consider a two-player, zero sum game with $S_1 = S_2 = \{1, \ldots 6\}$ and $\ell(i, j) = \mathbf{1}_{\{|i-j| \geq 2\}}$. Player 1 seeks to select i to minimize $\ell(i, j)$, i.e., to have $|i - j| \leq 1$, whereas player 2 seeks to select j to maximize $\ell(i, j)$, i.e. to have $|i - j| \geq 2$. Basically, player 1 is trying to get close to player 2 while player 2 wants to evade player 1.

(a) Show that for any Nash equilibrium, or equivalently, saddlepoint, (in mixed strategies), each player wins with equal probability.

Solution: By the theory of zero sum games, all saddle points, have the same value. The approach we follow is to identify a Nash equilibrium. (An alternative approach would be to identify the min max strategy for player 1 or the max min strategy for player 2.) Let p and q each be the uniform probability distribution over the set $\{2, 5\}$. We claim (p, q) is a Nash equilibrium. Indeed, given player 2 plays q, any pure strategy response i by player 1 is a best response, so p is also a best response for q. Similarly, given player one plays p, any pure strategy response j by player 2 is a best response so q is a best response to p. Thus, (p, q) is a Nash equilibrium. Each player wins with probability 1/2 for this, and therefore for any, Nash equilibrium.

(b) Find the minmax strategy for player 1 (it is unique) and all the maxmin strategies for player 2. Simplify your answer as much as possible.

Solution: A probability distribution $p = (p_1, \ldots, p_6)$ is a min max strategy for player one if and only if the maximum (over actions of player 2) loss of player 1 is equal to the

minimum possible value, already found to be 1/2. That is, if and only if:

$$\max\{p_3+p_4+p_5+p_6, p_4+p_5+p_6, p_1+p_5+p_6, p_1+p_2+p_6, p_1+p_2+p_3, p_1+p_2+p_3+p_4\} = 1/2$$

The probabilities p_2, p_3, p_4, p_5 each appear three times within the above sums and the p_1 and p_6 appear four times. So the average of the six sums within the braces is $(3 + p_1 + p_6)/6$. Since the maxium must be greater than or equal to the average, it follows that $p_1 = p_6 = 0$ for any minmax strategy of player 1. Then the average and the maximum of the six terms are both equal to 1/2. So all six terms in braces must equal 1/2 and $p_1 = p_6 = 0$. It follows that $p_2 = p_5 = 1/2$, and $p = (0, 1, 0, 0, 1, 0)$.
A probability distribution $q = (q_1, \ldots, q_6)$ is a max min strategy for player 2 if and only if the minimum (over actions of player 1) reward of player 2 is equal to the maximum possible value, already found to be 1/2. That is, if and only if:

$$\min\{q_3+q_4+q_5+q_6, q_4+q_5+q_6, q_1+q_5+q_6, q_1+q_2+q_6, q_1+q_2+q_3, q_1+q_2+q_3+q_4\} = 1/2$$

which, since the first and sixth sums are greater than or equal to some other sum, is equivalent to

$$\min\{q_4 + q_5 + q_6, q_1 + q_5 + q_6, q_1 + q_2 + q_6, q_1 + q_2 + q_3\} = 1/2$$

Since the terms $q_4+q_5+q_6$ and $q_1+q_2+q_3$ add to one, it is necessary that $q_4+q_5+q_6 = q_1+q_2+q_3$. In summary, the maxmin strategies of player 2 are all probability distributions q such that:

$$\begin{aligned} q_4 + q_5 + q_6 = q_1 + q_2 + q_3 &= 1/2 \\ q_1 + q_2 + q_6 &\geq 1/2 \\ q_1 + q_5 + q_6 &\geq 1/2 \end{aligned}$$

or

$$\begin{aligned} q_4 + q_5 + q_6 = q_1 + q_2 + q_3 &= 1/2 \\ 1/2 - q_3 + q_6 &\geq 1/2 \\ q_1 + 1/2 - q_4 &\geq 1/2 \end{aligned}$$

which reduces to the most simple form:

$$\begin{aligned} q_4 + q_5 + q_6 = q_1 + q_2 + q_3 &= 1/2 \\ q_3 \leq q_6 \text{ and } q_4 &\leq q_1 \end{aligned}$$

Intuitively speaking, player 2, who wants to evade player 1, needs to compensate for playing the strategies 3 or 4 near the middle by playing strategies on the ends. Some examples of maxmin strategies of player 2 are $(1/2, 0, 0, 0, 0, 1/2), (0, 0, 1/2, 0, 0, 1/2)$, and $(1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$.

7. **[Simulations for pursuit evasion game]**
Run Python simulations of repeated play of the game in the previous problem. See the Jupyter notebook at
http://nbviewer.jupyter.org/urls/courses.engr.illinois.edu/ece586GT/fa2017/ece586GTps3.ipynb?flush_cache=true. You are free to make use of the notebook for your assignment, for example by modifying some part of it. Investigate three scenarios, with both players using:

(a) Fictitious play. At each time t, each player plays a strategy that is a pure strategy best response to the empirical average of the previous $t-1$ plays of the other player. You can break ties any way you'd like, and (b)

(b) Weighted predictors with exponential weights $w_{i,t-1} = \mathrm{e}^{-\eta_t L_{i,t-1}}$ with $\eta_t = \sqrt{8(\ln(6)/t}$.

(c) Weighted predictors with exponential weights $w_{i,t-1} = \mathrm{e}^{-\eta_t L_{i,t-1}}$ with $\eta_t = \frac{\gamma}{t}$. (i.e. regularized fictitious play.)