

ECE 586GT: Problem Set 3

Fictitious play dynamics, minimum regret learning from forecasters

Due: Wednesday, October 14, 11:59 p.m.

Reading: Course notes, Chapters 3 and 4

1. [More on Dove-Hawk game]

Recall the Dove-Hawk game for two players:

	Dove	Hawk
Dove	4, 4	1, 5
Hawk	5, 1	0, 0

- Is this a potential game? If so, find the potential. If not, explain why.
- Identify the set of all correlated equilibria.
- Among the correlated equilibria, find one that maximizes the sum of the expected payoffs of the two players.

2. [A potential game for two players]

Consider the two player bi-matrix game:

	1	2	3
A	4, 4	1, 5	8, 9
B	5, 1	$x, 5$	$7, y$
C	$w, 5$	$7, 8$	$z, 5$

- For what values of w, x, y, z is the game a potential game and what is the potential function for that game?
- Identify all the pure Nash equilibria (using the values of w, x, y, z found in part (a)).

3. [On regularized fictitious play]

The purpose of this exercise is to fill in some details for the course notes Section 3.4. See that section for notation.

- Show that

$$\beta(r) \triangleq \arg \max_{p \in \Delta} \{p^T r + \tau H(p)\} = \sigma\left(\frac{r}{\tau}\right),$$

and the corresponding maximum value is $F(r) = \tau \ln \sum_a e^{r_a/\tau}$.

- Show that for a probability distribution q , $F(r) - q^T r - \tau H(q) = \tau D(q \| \beta(r))$. Hint: Calculate the righthand side using the expression for $D(\cdot \| \cdot)$ and β .
- Explain why Proposition 3.18 about convergence of stochastic approximation applies to the equation $\mu_1^{t+1} = \mu_1^t + \alpha_t(\beta(r) - \mu_1^t + D_{t+1})$.

4. [Hannan consistent player vs. arbitrary player]

Consider a finite two-player zero-sum game such that $\ell : S_1 \times S_2 \rightarrow \mathbb{R}$ 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.
- (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.

5. **[Blackwell approachability for a simple game with vector valued payoff]**

Consider a two-player game with strategy spaces

$$S_1 = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} \text{ and } S_2 = \{v \in \mathbb{R}^2 : \|v\| \leq 1\}$$

where $\|v\|$ is the Euclidian norm of v , 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 halfspace H in the plane can be expressed as $H = \{v : v \cdot u \leq c\}$, where u is a nonzero vector in $\mathbb{R}^2$ and $c \in \mathbb{R}$. Under what condition on u and c is this halfspace approachable (for player 1)?
- (b) Let $R(\delta)$ denote the square region $[-\delta, \delta]^2$ in the plane. Find the minimum value of δ so that $R(\delta)$ is approachable.

6. **[Exponential regret strategy vs. an oblivious player for rock, paper, scissors game]**

Consider two players playing the zero sum rock (R), paper (P), scissors (S) game, for $n = 1000$ rounds. Focus on the cumulative loss of player 1, in the case player 2 plays R for the first 500 rounds and then P for the remaining 500 rounds. (Player 2's strategy is far from being a stationary one.)

- (a) Describe how you think you would do against player 2, if all you knew to begin with is that you were to play 1000 times, and you saw the outcome after each play. (This is highly subjective—there is no single correct answer.)
- (b) Identify the cumulative loss functions $L_{R,t}, L_{P,t}, L_{S,t}$ for constant play R, P , or S by player one. (Note: For example, $L_{P,t} = -t$ for $0 \leq t \leq n/2$ because playing P wins each of the first 500 games, or equivalently, results in loss -1 in each of those games.)
- (c) Suppose player one plays the exponential weighted strategy: $\hat{p}_{t+1}(i) = \frac{e^{-\eta L_{i,t}}}{D_t}$ for $i \in \{R, P, S\}$, where $\eta = 0.1$ (which is near $\sqrt{8(\ln 3)/1000}$), and D_t is the normalizing constant making $\hat{p}_{t+1}$ a probability distribution. Show that the expected number of times player one wins is less than or equal to 501.

7. **[Strong law of large numbers for bounded martingale difference sequences]**

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 4.12.)