First type of question arising from Project 2: Reciprocity and the success of TFT strategies

Is there a pure NE for the IPD?

In Project 2, the Iterated Prisoner Dilemma game is considered. What is assumed that the two players choose a vector $\mathbf{p} = (p_1, \ldots, p_4)$ resp. $\mathbf{q} = (q_1, \ldots, q_4)$. Here, $p_i, q_i \in [0, 1]$ and $p_1, \ldots, p_4$ determine the probability of cooperation following previous play of the two players of $xy \in (cc, cd, dc, dd)$ (c=cooperation vs d=defection). Define the payoff vector

$$\mathbf{S}_X = (R, , S, T, P) \text{ and } \mathbf{S}_Y = (R, T, S, P).$$

Here R=reward, S=sucker, T=temptation , P=punishment, and a common choice is $(T, R, P, S) = (5, 3, 1, 0)$.

In Press-Dyson's paper it is shown that if we define

$$D(\mathbf{p}, \mathbf{q}, f) = \det \begin{pmatrix} p_1q_1 - 1 & p_1 - 1 & q_1 - 1 & f_1 \\ p_2q_3 & p_2 - 1 & q_3 & f_2 \\ p_3q_2 & p_3 & q_3 & f_3 \\ p_4q_4 & p_4 & q_4 & f_4 \end{pmatrix}$$

then the asymptotic expected score is

$$s_X = \mathbf{v} \cdot \mathbf{S}_X = \frac{\mathbf{v} \cdot \mathbf{S}_X}{v \cdot \mathbf{1}} = \frac{D(\mathbf{p}, \mathbf{q}, \mathbf{S}_X)}{D(\mathbf{p}, \mathbf{q}, \mathbf{1})} \tag{1}$$

$$s_Y = \mathbf{v} \cdot \mathbf{S}_Y = \frac{\mathbf{v} \cdot \mathbf{S}_Y}{v \cdot \mathbf{1}} = \frac{D(\mathbf{p}, \mathbf{q}, \mathbf{S}_Y)}{D(\mathbf{p}, \mathbf{q}, \mathbf{1})} \tag{2}$$

Here $\mathbf{v}$ is the stationary measure associated to the Markov chain

$$M = \begin{pmatrix} p_1q_1 & p_1(1 - q_1) & (1 - p_1)q_1 & (1 - p_1)(1 - q_1) \\ p_2q_3 & p_2(1 - q_3) & (1 - p_2)q_3 & (1 - p_2)(1 - q_1) \\ p_3q_2 & p_3(1 - q_2) & (1 - p_3)q_2 & (1 - p_3)(1 - q_2) \\ p_4q_4 & p_4(1 - q_4) & (1 - p_4)q_4 & (1 - p_4)(1 - q_4) \end{pmatrix},$$

i.e., the left eigenvector of M corresponding to eigenvalue 1. This matrix has a unique stationary vector $\mathbf{v}$ if and only if $D(\mathbf{p}, \mathbf{q}, \mathbf{1}) \neq 0$.

Remark 1. Note that even if $D(\mathbf{p}, \mathbf{q}, \mathbf{1}) = 0$ the expressions (1),(2) remain bounded, and so as $D(\mathbf{p}, \mathbf{q}, \mathbf{1}) \to 0$ there is a unique limit of these expressions. (There are quite a few, but perhaps only something like 20, choices for $\mathbf{p}, \mathbf{q}$ for which $D(\mathbf{p}, \mathbf{q}, \mathbf{1}) = 0$. The proof of a lemma from Wang-Lin's paper which shows that $D(\mathbf{p}, \mathbf{q}, \mathbf{1} \leq 0$ is helpful for this.)

Remark 2. $\mathbf{p}, \mathbf{q} \in [0, 1]^4$ can be viewed as pure actions. Mixed actions would correspond to probability measures in this space.

Question 1. Is there a pure Nash equilibrium for this infinite dimensional game, with payoffs determined by (1),(2) where we take the limit values.

Question 2. Work out precisely what happens when $D(\mathbf{p}, \mathbf{q}, \mathbf{1}) = 0$. I guess s_X, s_Y are then multivalued (depend on the initial choice of play).

Question 3. Is there a mixed Nash equilibrium for this infinite dimensional game, with payoffs determined by (1),(2).

Remark 3. Perhaps in Question 1 the existence can be proved abstractly, or perhaps can be found by a computer search.

Second type of question arising from Project 2: Reciprocity and the success of TFT strategies
The role of memory in IPD?

In Press-Dyson it is shown that if one player uses a one memory strategy then the other player has no benefit of using a longer term strategy.

Let us assume that an IPD tournament is played so that players do not know who they play in each round. So then perhaps this result is not so relevant.

Question 4. Does Press-Dyson's result apply if three players are involved? If not, show precisely in what way memory can be used.

Question 5. What is the role of memory in IPD tournaments. Can this be used to determine the environment you are up against?

Question 6. If you do choose a stationary (pure action) but either draw actions according to some probability measure on $[0,1]^4$. What improvements can one use? How is this related to memory?

In this project you are allowed to use books and the hand-outs, but it is important that you write up your project by yourself and that you can explain what you have written in detail during a short oral.

Attached to this project description are

- a section from the book of Sigmund (The calculus of selfishness),
- a paper by W.H. Press and F.J. Dyson `http://www.pnas.org/cgi/doi/10.1073/pnas.1206569109`.
- You are also asked to look at various discussion of the paper of Press and Dyson on the internet, e.g.

  `https://golem.ph.utexas.edu/category/2012/07/zerodeterminant_strategies_in.html`,

  `http://www.technologyreview.com/view/428920/the-emerging-revolution-in-game-theory/`

  `https://plato.stanford.edu/entries/prisoner-dilemma/#ZeroDeteStra`.
- A good paper discussing Press and Dyson's paper is A. Stewart and J.B. Plotkin, "Extortion and Cooperation in the Prisoner's Dilemma," Proceedings of the National Academy of Sciences, 109: 10134–10135: `http://www.pnas.org/cgi/doi/10.1073/pnas.1208087109`.

---

1. Submit your solutions of the selected exercises from the lecture notes (which are the same for all projects).
2. As discussed in the lectures, over the last 40 years every year tournaments were held, in which an interative version of the prisoner dilemma (or donation game) (IPR) was played. These tournaments are often called after its founder, Axelrod. In this project you are asked to explore the success of the TFT strategies in these tournaments. In a few paragraphs, explain the rules of the Axelrod tournaments, the various strategies that were used, and the outcomes in the tournament.
3. Discuss in a few pages what arguments are given in Sigmund's book, based on evolutionary game theory (and replicator dynamics), of the success of TFT games and explain why in these arguments the success of TFT depends on the environment in which it is used (i.e. on which other strategies are used).
4. Describe the results from the paper by Press and Dyson in detail.

5. Explain why the paper of Press and Dyson has created so much excitement, and discuss to what extend this excitement was warranted.

6. Run the Axelrod tournaments through one of the following

   - https://github.com/Axelrod-Python/Axelrod, see also https://vknight.org/unpeudemath/code/2015/02/20/an-iterated-prisoners-dilemma-on-github.html.

   - https://github.com/cristal-smac/ipd.

   Carefully document how you ran this code and your findings from these tournaments. Formulate your own strategy (for example some variant of TFT) and run this strategy in these tournaments. How does it do?

7. [**Mastery question for 4th year and MSc students**] Discuss the following two papers

   [-] C. Hilbe, M.A. Nowak, and K. Sigmund, 2013, "Evolution of extortion in Iterated Prisoner's Dilemma games," Proceedings of the National Academy of Sciences, 110 (17): 6913–6918. https://doi.org/10.1073/pnas.1214834110.

   [-] S. Wang and F. Lin, 2020, "Nice Invincible Strategy for the Average-Payoff IPD", https://aaai.org/ojs/index.php/AAAI/article/view/5604

   There are quite a few other recent papers on this topic, and you are also allowed to discuss any of those (instead), provided your discussion goes into sufficient depth:

   [-] E. Akin, 2013, "The Iterated Prisoner's Dilemma: Good Strategies and Their Dynamics," https://arxiv.org/abs/1211.0969.

   [-] P. Mathieu and J.P. Delahaye, 2017, "New Winning Strategies for the Iterated Prisoner's Dilemma", Journal of Artificial Societies and Social Simulation 20 (4) 12, http://jasss.soc.surrey.ac.uk/20/4/12.html.

   [-] Y. Murase, S.K. Baek, 2020: "Five rules for friendly rivalry in direct reciprocity", https://arxiv.org/abs/2004.00261

   [-] M. Harper, V. Knight et al., 2017: "Reinforcement learning produces dominant strategies for the Iterated Prisoner's Dilemma" Plos One 12 (12) , e0188046. https://doi.org/10.1371/journal.pone.0188046