M?PA48 Games and Dynamics: Lecture Notes

Sebastian van Strien (Imperial College)

Autumn 2015

Contents

0 Introduction

0.1 Practial Arrangement

- The lectures for this module will take place **Monday 9-11, Thursday 1-2**.
- This course does not require any background in game theory, and in fact the overlap with the game theory course that is offered by the department is minimal.
- The way this course will be examined, will be agreed in week 3. It depends on how many students take this course, on whether there is a preference to do projects as part of the examination.
- **Questions are most welcome**, during or after lectures and during office hour.
- My office hour is to be agreed with students reps. Office hour will in my office 6M36 Huxley Building.

0.2 What is this course about?

The notion of Nash equilibrium is prevalent in many areas of science: economics, biology, engineering etc. After all, many phenomena can be described as versions where someone, or something, tries to optimise.

The aim of this course is to highlight some situations where the notion of Nash equilibrium, or related notions, are given a more dynamic interpretation. So a Nash equilibrium would the stationary point of some differential equation, or of some other dynamical process.

Example 1. Consider a population of birds where some will always fight about a grain (let us call such a bird a hawk), and others will always do some posturing but will never fight (doves). The payoff of getting the grain is G, and the price for getting hurt is $-C$. We assume that $0 < G < C$. If a hawk meets a hawk, it either wins (and gets payoff G) or looses (and get payoff $-C$). On average this means the payoff is $(G + -C)/2$. In this way we get the payoff matrix

$$\begin{array}{r|cc} \text{meeting} & \text{Hawk} & \text{Dove} \\ \hline \text{payoff to Hawk} & \frac{G-C}{2} & G \\ \text{payoff to Dove} & 0 & \frac{G}{2} \end{array}.$$

How is it that not the entire species develops hawk-behaviour? Suppose that the frequency of hawks in the population is x and doves is $1 - x$. Then the average increase of 'fitness' is

$$\begin{array}{ll} x(G-C)/2 + (1-x)G & \text{for hawks} \\ x \cdot 0 + (1-x)G/2 & \text{for doves.} \end{array}$$

If $x = 1$ then the increase of fitness of hawks < 0 and of doves > 0, and so the number of doves will increase and the number of hawks will decrease. (Hawks are constantly fighting and getting injured, whereas the doves will occasionally get lucky.) When $x = 0$, the fitness is G resp. $G/2$, so the number of hawks will increase. Equality holds when $x = G/C$.

Example 2 (Prisoner dilemma)**.** Consider two prisoners, which are in separate rooms so that they cannot communicate. The prisoners get a higher reward by betraying the other (defecting), but if both coorporate (so stay silent) they get a reduced sentence. For example we may have the following situation:

$$\begin{array}{r|cc} \text{Prisoner II} & \text{Coop} & \text{Defects} \\ \hline \text{Pris. I Coop} & -1,-1 & -3,0 \\ \text{Pris I Defects} & 0,-3 & -2,-2 \end{array}.$$

This table describes the payoff (the number of years prison sentence) in various scenarios. For example if prisoner II defects but prisoner I cooperates, then prisoner II will be released and prisoner I will be 3 years in prison. What should the prisoners do? If II cooperates then I is better off to defect (he then gets 0 years rather than 1 year prison sentence). If II defects then he still better to defect (he gets 2 years rather than rather than 3 years). The same holds for II. So the rational behaviour is for both prisoners is to defect, resulting in a prison sentence of two years for each.

Example 3 (Repeated prisoner dilemma)**.** Suppose that the previous set-up is repeated every year? Or to discuss a slight variant, where two players are asked every week to make a donation of £5. If so, the other player gets a donation of £15, otherwise nothing. So the situation is described by

$$\begin{array}{lcc} \text{II} & \text{donates} & \text{declines} \\ \text{I donates} & 10,10 & -5,15 \\ \text{I declines} & 15,-5 & 0,0 \end{array}.$$

Of course in one-step this is again a prisoner dilemma. If this play is repeated many times then the considerations of the players change of course. We will discuss this situation in this course. (A political scientist called Axelrod, even organises computer tournaments which explore which strategy is the most optimal. One strategy is called TFT (Tit for Tat).)

Example 4. Different types of game dynamics In this course we will consider various types of game dynamics. For example, the first section we will consider the well-known replicator dynamics.

To emphasise that it is important to consider the detailed set-up of the game, let us consider the following:

Example 5 (Parrondo paradox)**.** Consider two games Game A and Game B:

- In Game A, you lose £1 every time you play.
- In Game B, you count how much money you have left. If it is an even number, you win £3. Otherwise you lose £5.

Say you begin with £100. If you start playing Game A exclusively, you will obviously lose all your money in 100 rounds. Similarly, if you decide to play Game B exclusively, you will also lose all your money in 100 rounds.

However, consider playing the games alternatively, starting with Game B, followed by A, then by B, and so on (BABABA...). It should be easy to see that you will steadily earn a total of £2 for every two games.

Thus, even though each game is a losing proposition if played alone, because the results of Game B are affected by Game A, the sequence in which the games are played can affect how often Game B earns you money, and subsequently the result is different from the case where either game is played by itself.

Example 6 (Learning algorithms: Reinforcement learning)**.** The notion of payoff to players also leads to various learning principles: the higher the payoff from a given action is, the more likely this action will be taken in the future. There are various models which make this intuitive notion precise.

Example 7 (Learning algorithms: No-regret learning)**.** A different variant of a learning algorithm is that of no-regret learning. This is based on the idea that if a different action in the past would have given a better payoff, assuming the other player would have done the same.

0.3 References for the various chapters

- Chapter 1: mainly the book of Hofbauer & Sigmund, *Evolutionary games and population dynamics* but another useful book to consult is Weibull, *Evolutionary game theory*.

- Chapter 2: chapter 3 of Sigmund's book *The Calculus of Selfishness*.

- Chapter 3: more on this can be found in Hofbauer, *Deterministic Evolutionary Game Dynamics*, Proceedings of Symposia in Applied Mathematics Volume 69, 2011.

- Chapter 4: more on the classification of replicator dynamics can be found in Hofbauer & Sigmund, *Evolutionary games and population dynamics* but a more detailed description can be found in chapter 3 of Cressman, *Evolutionary Dynamics and Extensive Form Games*.

  The description of a chaotic replicator dynamics system is given in Satoa, Akiyamab and Crutchfield, *Stability and diversity in collective adaptation*, Physica D, 210, 2015, 21-57.

  For results on chaotic best response dynamics, see for example my papers on game theory: `http://wwwf.imperial.ac.uk/~svanstri/publications_by_subject.php`

- Chapter 5 follows Ostrovski & van Strien, *Payoff performance of fictious play, Journal of Dynamics and Games*, vol 1, issue 4, October 2014

- Chapter 6 follows essentially Posch, *Cycling in a stochastic learning algorithm for normal form games*, J Evol Econ (1997) 7: 193-207. But there is an extensive literature on this.

Some of this work is grounded in the field of behavioural economics, so to model how people learn, e.g. Erev & Roth, *Predicting how people play games: Reinforcement learning in experimental games with unique*, mixed strategy equilibria. Amer. Econ. 1998, Rev. 88, 848?881.

For a discussion on approximating discrete 'random' dynamical systems by differential equations can be found in for example Benaïm, *Dynamics of stochastic approximation algorithms*, in: Seminaire de Probabilité,XXXIII, Lecture Notes in Mathematics, vol. 1709, Springer, Berlin, 1999, pp. 1-68.

For a starting point on reinforcement learning work in the machine learning community see `https://en.wikipedia.org/wiki/Reinforcement_learning` and Sutton, *Reinforcement Learning: An Introduction*, 1998 or Murphy, *Machine Learning: A Probabilistic Perspective*, 2010.

- Chapter 7 discusses Hart & Mas-Colell, *A simple adaptive procedure leading to correlated equilibrium*, Econometrica, Vol. 68, No. 5 September, 2000, 1127-1150. Hart, *Adaptive heuristics*, Econometrica, Vol. 73, No. 5 September, 2005, 1401-1430 has a more extensive discussion of the literature.

  For regret dynamics from a computer science (machine learning) point of view see for example `http://theory.stanford.edu/~tim/f13/l/l18.pdf` and `http://theory.stanford.edu/~tim/f13/l/l17.pdf`

- More general references for the chapters on learning are: Fudenberg & Levine, *The Theory of Learning in Games*. MIT Press. (1999) and Young, *Strategic learning and Its limits*,.Oxford, U.K, (2004), or from the machine learning point of view, see for example Nisan, Roughgarden, Tardos and Vazirani, *Algorithmic Game Theory*, 2007.

1 Replicator dynamics for one population

1.1 Nash equilibrium of one population

We consider a large population where each individual can have one of a finite set of pure strategies $\{1, ..., n\}$. You should think of these as individuals which can have one of n different traits (e.g. colour of eyes, fighting behaviour, personally characteristics etc). Let x_i denote the frequency of strategy i. So $(x_1, \dots, x_n)$ is a probability vector.

Let $\Delta_n = \{x \in \mathbb{R}^n; 0 \le x_i \le 1, x_1 + \dots + x_n = 1\}$ be the $(n-1)$-dimensional simplex. So $(x_1, \dots, x_n) \in \Delta_n$. Usually we will fix n and write Δ.

The **payoff** to strategy i in a population x is $a_i(x)$, with $a_i : \Delta \to \mathbb{R}$ a continuous function (population game).

Let us for the moment consider the case of a symmetric two person game with $n \times n$ payoff matrices $A = (a_{ij})$ and $B = A^{tr}$. In the current context this means that you have two populations I and II which compete. The payoff of population I depends on (i.e. 'success') of a trait i depends on how it does when it competes with another trait j. With random matching (that is, random encounters) this leads to the following **linear payoff function of population I**

$$a_i(x) = \sum_j a_{ij} x_j = (Ax)_i.$$

If a population I with frequency distribution y encounters another population II with frequency distribution x then the resulting payoff for the population with distribution y will be

$$\text{Payoff}_{\text{I}}(y, x) := y \cdot Ax.$$

That this is a **symmetric two person** means that we assume that in this situation, population II receives an equal payoff of

$y \cdot Ax$. Since

$$y \cdot Ax = y^{tr} Ax = x^{tr} A^{tr} y$$

the payoff for player II is described by the matrix $B = A^{tr}$ and

$$\text{Payoff}_{\mathbf{II}}(x, y) = x \cdot By \text{ with } B = A^{tr}.$$

We say that $\hat{x} \in \Delta$ is a **Nash equilibrium (NE)** iff

$$x \cdot A\hat{x} \leq \hat{x} \cdot A\hat{x}, \forall x \in \Delta.$$

and a **strict Nash equilibrium** if

$$x \cdot A\hat{x} < \hat{x} \cdot A\hat{x}, \forall x \in \Delta \text{ with } x \neq \hat{x}.$$

Note that $x \cdot A\hat{x} \leq \hat{x} \cdot A\hat{x}, \forall x \in \Delta$ means that strategy $\hat{x}$ cannot be 'improved' by another strategy x.

An equivalent way of formulating the notion of Nash equilibrium is to define the **best response** map

$$\mathcal{B}R(x) = \arg\max_{y \in \Delta} y \cdot Ax.$$

Then $\hat{x}$ is a NE iff $\hat{x} \in \mathcal{B}R(\hat{x})$.

Example 8. Consider a game determined by $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. What are its Nash equilibria? To see this, note that

$$\mathcal{B}R(x) = \begin{cases} e_1 & \text{if } x_1 > x_2 \\ e_2 & \text{if } x_1 < x_2 \\ \Sigma & \text{if } x_1 = x_2 \end{cases}$$

So $e_i \in \mathcal{B}R(e_i)$ and $\mathcal{B}R \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \ni \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$. So e_1, e_2 and $z := (e_1 + e_2)/2 := \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$ are the Nash equilibria. Note that $Az = z$ and so $x \cdot Az = 1/2$ for **each** $x \in \Sigma$. So z is not a strict NE. On the other hand, e_1, e_2 are both strict NE.

Example 9. Consider a game determined by $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. What are its Nash equilibria? Note that Δ in this case is a triangle. $\mathcal{B}R$ takes values e_1, e_2, e_3 on three convex regions (draw this) which meet at $(1/3, 1/3, 1/3)$. At this midpoint, one has that $(1/3, 1/3, 1/3) \in \mathcal{B}R(1/3, 1/3, 1/3) = \Sigma$ and so this is a NE. Along the line segment where the e_1 and e_2 meet, $\mathcal{B}R =< e_1, e_2 >$ and so where this line meets Σ we get another NE. Continuing this analysis, we see there are 7 NE's.

1.2 Evolutionary stable strategies

$\hat{x}$ is an **evolutionary stable equilibrium (ESS)** if for all $x \in \Delta, x \neq \hat{x}$ one has for $\epsilon > 0$ small enough,

$$x \cdot A(\epsilon x + (1 - \epsilon)\hat{x}) < \hat{x} \cdot A(\epsilon x + (1 - \epsilon)\hat{x}). \qquad (1)$$

Here the size of $\epsilon > 0$ is allowed to depend on x.

Lemma 1. strict NE $\implies$ ESS $\implies$ NE.

Remark: Every game has a Nash equilibrium, but there are games without an ESS.

Proof. First assume $\hat{x}$ is a strict NE. Then $x \cdot A\hat{x} < \hat{x}A\hat{x}$. This inequality is what the ESS condition (1) reduces to if we take $\epsilon = 0$. By continuity the ESS condition then also holds for $\epsilon > 0$ small.

Now assume that x is an ESS. For each $x \neq \hat{x}$ we can let $\epsilon \to 0$ in the ESS condition and we obtain $x \cdot A\hat{x} \leq \hat{x}A\hat{x}$. □

Example 10. Consider a game determined by $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. What are its ESS? We know this matrix has three NE's. Let us check which of these are ESS's. Let us check whether $z = (e_1 + e_2)/2$ is an ESS. Note that $x \cdot Az = 1/2$ for each x. For

z to be an ESS, we need that for $\epsilon > 0$ small, and all $x \neq z$ we have $x \cdot A(\epsilon x + (1-\epsilon)z) < z \cdot A(\epsilon x + (1-\epsilon)z)$. This reduces to $x \cdot Ax < z \cdot Ax$. This is supposed to hold for all $x \neq z$ so in particular for $x = e_1$, but this is clearly not true. So z is not an ESS. Note that, in fact, we could have used the next lemma to conclude that z is not an ESS.

Let us now show that e_1 is an ESS. So we need to show that for all $x = (x_1, x_2) \neq \hat{e}_1$ and all $\epsilon > 0$ sufficiently small, $x \cdot A(\epsilon x + (1-\epsilon)e_1) < e_1 \cdot A(\epsilon x + (1-\epsilon)e_1)$ when $x \neq e_1$. This is equivalent to $\epsilon(x_1^2 + x_2^2) + (1-\epsilon)x_1 < \epsilon x_1 + (1-\epsilon)$ which holds since $x_1 \neq 1$. [If we take $\epsilon = 0$, then the ESS inequality becomes $x \cdot e_1 < e_1 \cdot e_1$ which clearly holds when $x \neq e_1$. So for $\epsilon > 0$ the ESS inequality also holds.] In the same way we get that e_2 is also an ESS.

Lemma 2. If $\hat{x}$ is a Nash equiliibrium then there exists $c \in \mathbb{R}$ so that $(A\hat{x})_i = c$ for each i with $\hat{x}_i > 0$. In particular, if $\hat{x} \in \operatorname{int}\Delta$ is a NE then there exists $c \in \mathbb{R}$ with $(A\hat{x})_i = c$ for each i. If $\hat{x} \in \operatorname{int}\Delta$ is an ESS, then there exists no other NE.

Proof. Substituting e_i for p in the definition of the NE, we get

$$e_i \cdot A\hat{x} \leq \hat{x} \cdot A\hat{x}.$$

This holds for all $i = 1, \dots, n$. Write $\hat{x} = \sum \lambda_i e_i$ with $\lambda_i \geq 0$ and $\sum \lambda_i = 1$. Summing over the previous inequality we get

$$\hat{x}A\hat{x} = \sum \lambda_i e_i \cdot A\hat{x} \leq \sum \lambda_i \hat{x} \cdot A\hat{x} = \hat{x} \cdot A\hat{x}.$$

But we would get strict inequality if $e_i \cdot A\hat{x} < \hat{x} \cdot A\hat{x}$ for some i for which $\lambda_i > 0$, which is clearly impossible. Hence $e_i \cdot A\hat{x} = \hat{x} \cdot A\hat{x}$ for all $i = 1, \dots, n$ for which $\hat{x}_i > 0$.

Hence, by what we used proved, if $\hat{x}$ is an interior NE, then for each $x \in \Delta$ one has $x \cdot A\hat{x} = c$ (here we use that x is a probability vector). Assume that $\hat{x}$ is also an ESS, i.e. for each $x \neq \hat{x}$ and $\epsilon > 0$ small, one has $x \cdot A(\epsilon x + (1-\epsilon)\hat{x}) <$

$\hat{x} \cdot A(\epsilon x + (1-\epsilon)\hat{x})$. But since $x \cdot A\hat{x} = \hat{x} \cdot A\hat{x} = c$, this inequality reduces to $x \cdot Ax < \hat{x} \cdot Ax$ for each $x \neq \hat{x}$. So $x \neq \hat{x}$ cannot be a NE. □

Lemma 3. The ESS assumption (1) is equivalent to the assumption that for all $y \neq \hat{x}$ sufficiently close to $\hat{x}$,

$$y \cdot Ay < \hat{x} \cdot Ay \tag{2}$$

Moreover, if $\hat{x} \in \operatorname{int}\Delta$ is an ESS then

$$y \cdot Ay < \hat{x} \cdot Ay \text{ for all } y \in \Delta \tag{3}$$

Proof. Each y which is close to $\hat{x}$ can be written in the form $y = \epsilon x + (1-\epsilon)\hat{x}$ where we choose $x \in \partial\Delta$ and so that moreover, if $\hat{x} \in \partial\Delta$ then x is chosen so that it is not in the same face of $\partial\Delta$ as $\hat{x}$. Since the points x of this type are not close to $\hat{x}$, we can find some $\epsilon_0 > 0$ so that for all such x the inequality (1) holds for $\epsilon = \epsilon_0$. Substituting the definiton of y, in (1) gives $x \cdot Ay < \hat{x} \cdot Ay$. Multiplying this inequality by ϵ and adding to both sides the inequality the term $(1-\epsilon)\hat{x} \cdot Ay$, gives the required inequality $y \cdot Ay < \hat{x}Ay$.

Now assume that $\hat{x} \in \operatorname{int}\Delta$ is an ESS. Then by the previous lemma, there exists c so that for all x, $x \cdot A\hat{x} = c = \hat{x} \cdot A\hat{x}$. This means that $x \cdot A(\epsilon x + (1-\epsilon)\hat{x}) < \hat{x} \cdot A(\epsilon x + (1-\epsilon)\hat{x})$ reduces to the required inequality $x \cdot Ax < \hat{x} \cdot Ax$ for all $x \in \Delta$. □

Example 11. Consider $A = \begin{pmatrix} \frac{G-C}{2} & G \\ 0 & \frac{G}{2} \end{pmatrix}$. Then $\begin{pmatrix} G/C \\ (C_G)/2 \end{pmatrix}$ is its unique ESS.

Example 12. Show that $A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$ does not have any ESS. What are its NE? Assume $x \in \operatorname{int}\Delta$ is a NE. Then there exists c so that $(Ax)_i = c$ for each i. So $x_1 - x_3 = -x_1 + x_3 = x_1 - x_1 = c$, which gives $c = 0$ and $x_i = 1/3$, and the point $(1/3, 1/3, 1/3)$ is a NE. Now consider whether x can be a NE when $x_3 = 0$. Then $Ax = (x_2, -x_1, x_1 - x_2)$. [Note I don't always write a column vector when I should.] In fact, one draw the set where $e_i \in \mathcal{B}R(x)$ consists of a convex region containing $(1/3, 1/3, 1/3)$ (see the lectures for a drawing). From this diagram follows that $(1/3, 1/3, 1/3)$ is the only NE. Is $\hat{x} = (1/3, 1/3, 1/3)$ a ESS? Again we need to consider the inequality $x \cdot A(\epsilon x + (1-\epsilon)\hat{x}) < \hat{x} \cdot A(\epsilon x + (1-\epsilon)\hat{x})$. This reduces to $x \cdot Ax < \hat{x} \cdot Ax$. That is, $x_1(x_2 - x_3) + x_2(-x_1 + x_3) + x_3(x_1 - x_2) < (1/3)[(x_2 - x_3) + (-x_1 + x_3) + (x_1 - x_2)]$. Note that both the left and right hand side are zero, so this the inequality does NOT hold.

1.3 Replicator dynamics

One proposal to describe a mechanism which explains why Nash equilibria and ESS can appear as a dynamic process is the following system of differential equations

$$\dot{x}_i = x_i((Ax)_i - x \cdot Ax), i = 1, \ldots, n. \tag{4}$$

Note that this implies that

$$\frac{d}{dt}\frac{x_i}{x_j} = \frac{x_i}{x_j}((Ax)_i - (Ax)_j). \tag{5}$$

Lemma 4 (Nash equilibria and equilibria of the ODE)**.** .

1. Any Nash equilibrium $\hat{x}$ is an equilibrium of this equation.
2. If $\hat{x}$ is the omega-limit of an orbit $x(t)$, and $\hat{x} \in \operatorname{int}\Delta$ then $\hat{x}$ is a NE.
3. If $\hat{x}$ is Lyapounov stable, then it is a NE.

Proof. By the previous lemma, if $\hat{x}$ is a Nash euqilibrium then there exists a c so that $(A\hat{x})_i = c$ for each i for which $\hat{x}_i > 0$. It follows that $(A\hat{x})_i - \hat{x} \cdot A\hat{x} = 0$ for each of these i. For the other i one has $\hat{x}_i = 0$.

If $\hat{x}$ is not a Nash equlibrium then there exists x so that $x \cdot A\hat{x} > \hat{x} \cdot A\hat{x}$. It follows that there exists i so that $e_i \cdot A\hat{x} > \hat{x} \cdot A\hat{x}$. Hence there exists $\epsilon > 0$ so that for x close to $\hat{x}$ (here we reuse the name x), $e_i \cdot Ax > x \cdot Ax > \epsilon$. Hence $\dot{x}_i > \epsilon x_i$ when x is close to $\hat{x}$ and so it is impossible that $x(t) \to \hat{x}$ as $t \to \infty$. □

Example 13. Give an example of a system for which not every stationary point $\hat{x}$ is a NE. (Hint: there may be indices i with $\hat{x}_i = 0$ when $(A\hat{x})_i > c$ where c is as above.)

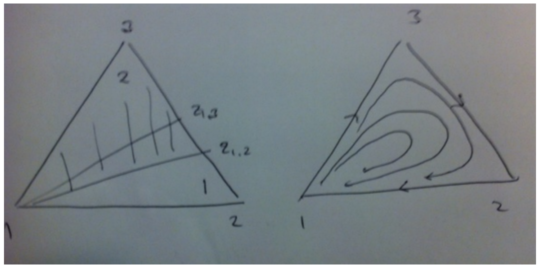
Example 14. Describe what happens for

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \tag{6}$$

What are its NE's? What are its ESS's? Let us first check whether A has any interior NE x. Then $(Ax)_i = c$ and so $x_2 = 2x_3 = x_3 = c$. So $c = x_3 = x_2 = 0$. So there is no interior NE. Looking at the level sets of $\mathcal{B}R$ we see that e_1 is the only NE. Is $\hat{x} = e_1$ a EES? Note that $A\hat{x} = 0$, so $x \cdot A(\epsilon x + (1-\epsilon)\hat{x}) < \hat{x} \cdot A(\epsilon x + (1-\epsilon)\hat{x})$ reduces to $x \cdot Ax < \hat{x} \cdot Ax$. This is equivalent to $x_1 x_2 + x_2 2x_3 + x_3 x_3 < x_2$ which is not the case when $x_1 = x_2 = 0, x_3 = 1$.

Solution example 14. Denote by $Z_{i,j}$ the set where the best response is indifferent between i, j. Then $Z_{1,2} = \{x_2 = 2x_3\}$, $Z_{1,3} = \{x_2 = x_3\}$ and $Z_{2,3} = \{2x_3 = x_3\} = \{x_3 = 0\}$. These are all lines through e_1. By considering the position, of these lines, it follows the triangle has regions where $\mathcal{B}R = 1$ and $\mathcal{B}R = 2$ Note that $Ax = \begin{pmatrix} x_2 \\ 2x_3 \\ x_3 \end{pmatrix}$, Hence $(x_2/x_3)' =$

$(x_2/x_3)(2x_3 - x_3)$, $(x_1/x_2)' = (x_1/x_2)(x_2 - 2x_3)$, $(x_1/x_3)' = (x_1/x_3)(x_2 - x_3)$. This shows that on the sides of the triangle there are no additional singularities. It follows that the phase portrait is as follows. Note that each of the corners e_i is a singularity, but only e_1 is a Nash equilibrium.



Theorem 1. If $\hat{x}$ is an ESS then it is asymptotically stable for the replicator system.

If $\hat{x} \in \operatorname{int}\Delta$ is an ESS then it globally attracts all initial points $x \in \operatorname{int}\Delta$.

Proof. Consider the function $P(x) = \prod x_i^{\hat{x}_i}$. Let us show that this has a unique maximum at $\hat{x}$. First notice that when f is a convex function on some interval I, then $f(\sum p_i x_i) \leq \sum p_i f(x_i)$ for $x_1, \dots, x_n \in I$ and all p_i with $p_i \geq 0$ and $\sum p_i = 1$. If f is strictly convex, then a strict inequality holds except when all the x_i are equal. Applying this to $f = \log$ on $[0, \infty]$ (which is concave, so we get the opposite inequality) gives $\sum \hat{x}_i \log(\frac{x_i}{\hat{x}_i}) = \sum_{\hat{x}_i > 0} \hat{x}_i \log(\frac{x_i}{\hat{x}_i}) \leq \log \sum_{\hat{x}_i > 0} x_i \leq \log \sum x_i = \log 1 = 0$. Hence $\sum_i \hat{x}_i \log x_i \leq \sum_i \hat{x}_i \log \hat{x}_i$ and so $P(x) \leq P(\hat{x})$ with inequality only if $x = \hat{x}$.

So let us now show that we can consider P as a Lyapounov function:

$$\frac{\dot{P}}{P} = \frac{d}{dt}(\log P) = \frac{d}{dt} \sum \hat{x}_i \log x_i = \sum_{\hat{x}_i > 0} \hat{x}_i \frac{\dot{x}_i}{x_i} =$$

$$= \sum \hat{x}_i((Ax)_i - x \cdot Ax) = \hat{x} \cdot Ax - x \cdot Ax$$

Since $\hat{x}$ is evolutionary stable, the equation (2) gives that the r.h.s. is > 0 and so $\dot{P} > 0$ for all $x \neq \hat{x}$ close to $\hat{x}$. It follows that orbits starting near $\hat{x}$ converge to $\hat{x}$.

If $\hat{x} \in \operatorname{int} \Delta$ then (3) implies that $\dot{P}/P > 0$ everywhere and so $\hat{x}$ attracts all points in $\operatorname{int} \Delta$. $\square$

Example 15. Consider the matrix $A = \begin{pmatrix} 0 & 6 & -4 \\ -3 & 0 & 5 \\ -1 & 3 & 0 \end{pmatrix}$. Show that $E = (1/3, 1/3, 1/3)$ is a rest point which is asymptotically stable. To see this, compute the eigenvalues of the linearisation at this fixed point. Show that this point is not an ESS, by showing that $e_1 = (1, 0, 0)$ is an ESS.

Solution example 15. Note that the lines $Z_{i,j}$ all go through E. $Ax = \begin{pmatrix} 6x_2 - 4x_3 \\ -3x_1 + 5x_3 \\ -1x_1 + 3x_2 \end{pmatrix}$. From this one can see that the lines $Z_{i,j}$ are as in the figure, and so E is a Nash equilibrium. This also determines the singularities and the arrows on the sides of the triangle, as $(x_i/x_j)' = (x_i - x_j)[(Ax)_i - (Ax)_j]$. Indeed, $Z_{2,3} \cap [e_2, e_3]$ and $Z_{1,2} \cap [e_1, e_2]$ are singularities, and of course e_1, e_2, e_3 are also singularities. Note that 2 and 3 are suboptimal strategies at $Z_{2,3} \cap [e_2, e_3]$ and so this point is not a Nash equilibrium. Similarly, e_2, e_3 are not Nash equilibria. On the other hand, $Z_{1,3} \cap [e_1, e_3]$ and e_1 are Nash equilibria. In summary, this game has three Nash equilibria and three additional singularities.

The singularity $Z_{2,3} \cap [e_2, e_3] = (0, 5/8, 3/8)$ is a saddle point. Indeed on $[e_2 e_3]$ we have $(x_2/x_3)' = (x_2/x_3)[5x_3 - 3x_2]$. This shows that the arrows along this side point towards $(0, 5/8, 3/8)$. In the assignments you are asked to show that this point is indeed a saddle point.

To compute the eigenvalues in $E = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$ we write $x_i = E + h_i$ and let h be the vector with components h_i. Note that $\sum h_i = 0$. Since all components of AE are equal we have that $h \cdot AE = 0$ and so

$$\begin{aligned} x \cdot Ax &= (E+h) \cdot A(E+h) \\ &= E \cdot AE + h \cdot AE + E \cdot Ah + O(h^2) \\ &= E \cdot AE + E \cdot Ah + O(h^2). \end{aligned} \tag{7}$$

and

$$(Ax)_i - x \cdot Ax \quad = (Ah)_i - E \cdot Ah + O(h^2). \tag{8}$$

Taking $\mathbb{1}$ to be the vector with 1's we get

$$\begin{aligned} E \cdot Ah &= (1/3)\mathbb{1} \cdot Ah \\ &= (1/3)(-4h_1 + 9h_2 + h_3) \\ &= (-5/3)h_1 + (8/3)h_2 \end{aligned}$$

and

$$\begin{aligned} (Ah)_1 &= 6h_2 - 4h_3 = 4h_1 + 10h_2, \\ (Ah)_2 &= -3h_1 + 5h_3 = -8h_1 - 5h_2. \end{aligned}$$

So $\dot{x}_i = x_i((Ax)_i - x \cdot Ax)$ gives

$$\begin{aligned} \dot{h}_1 &= ((1/3) + h_1)\left((17/3)h_1 + (22/3)h_2 + O(h^2)\right)) = \\ &= (1/9)\left(17h_1 + 22h_2\right) + O(h^2). \\ \dot{h}_2 &= (1/9)\left(-19h_1 - 23h_2\right) + O(h^2). \end{aligned}$$

So the linear part is equal to

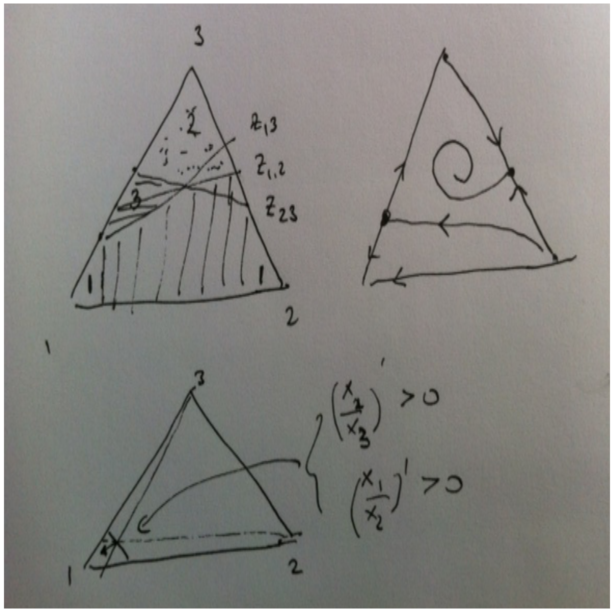
$$(1/9)\begin{pmatrix} 17 & 22 \\ -19 & -23 \end{pmatrix}$$

The eigenvalues of the matrix are $(1/3)(-1 \pm i\sqrt{2})$.

In Assignment 1, you are asked to check and correct this calculation, including that of the linear part and eigenvalues.

To see that e_1 is an ESS it is sufficient to check that $(x - e_1) \cdot A(\epsilon x + (1-\epsilon)e_1) < 0$ when $\epsilon > 0$ small. Another way of seeing this, is to observe that it is sufficient to show that $P = x_1$ is a strict Lyapounov function. (To see that this is sufficient, have a look at the proof of the previous theorem. There it is shown that $\dot{P}/P = \hat{x} \cdot Ax - x \cdot Ax$ and by Lemma 3 ESS is equivalent to the statement that this term is positive for x close to $\hat{x}$.) But we have that $(x_1/x_3)' = (x_1/x_3)[(Ax)_1-(Ax)_3]$ and $(x_1/x_2)' = (x_1/x_2)[(Ax)_1 - (Ax)_2]$ where the square bracket terms are both positive. This means that the speed vector along the line $P = x_1 = \epsilon$ lies in the cone in the figure, and so P is strictly increasing.

Additional arguments are needed to show that the saddle-separatrices are as shown.


3
$\hat{x}$
2
Z_{13}
$Z_{1,2}$
Z_{23}
3
1
1
2
1
3
$\left(\frac{x_1}{x_3}\right)' > 0$
$\left(\frac{x_1}{x_2}\right)' > 0$
ϵ
1
2

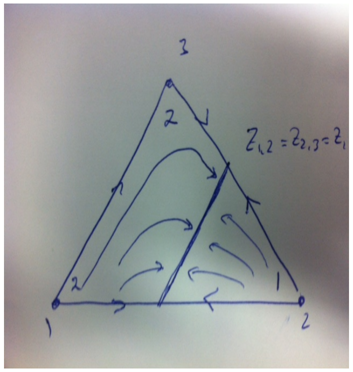
In the current setting, we say that A corresponds to a zero-sum game if $a_{ij} = -a_{ji}$. In this case $x \cdot Ax = -x \cdot Ax = 0$ and the replicator dynamics becomes

$$\dot{x}_i = x_i(Ax)_i.$$

Example 16. Consider the following matrices, determine the corresponding NE's and the phase diagrams of the replicator dynamics. $A = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$; $Z_{1,2}$ and $Z_{1,3}$ correspond to $2x_2 = 2x_1 + 2x_3$ resp $2x_1 + 2x_3 = x_1 + x_2 + x_3$. These are the same lines. $Z_{1,3}$ corresponds to $2x_2 = x_1 + x_2 + x_3 = 1$, so $x_2 = 1/2$. So $Z_{1,3} = Z_{1,2} = Z_{2,3}$ and this lines consists entirely of NE's. These are stationary points of the system, so in particular there are no ESS's. In summary, this system has infinitely many NE's and three additional singularities.

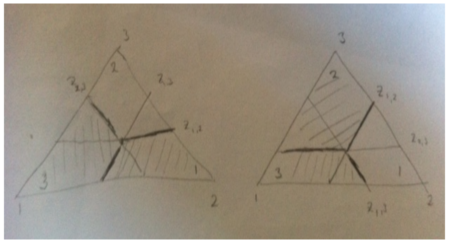
The arrows along the boundary can be seen by using $(x_i/x_j)' = (x_i/x_j)[(Ax)_i - (Ax)_j]$. Along $[e_1, e_2]$ we get $(Ax)_1 - (Ax)_2 = (2x_2 - 2x_1)$, so a sign change at $x_1 = x_2 = 1/2$. Along $[e_1, e_3]$ we get $(Ax)_1 - (Ax)_3 = (2x_2 - 1) = -1 < 0$ and along $[e_2, e_3]$ we get $(Ax)_2 - (Ax)_3 = (2x_1 + 2x_3 - 1) = 2x_3 - 1$ which has a sign change. Along $Z_{1,2} = Z_{1,3}$ we have that $Ax = (1, 1, 1)$ so this means that all these points are singularities of the replicator system. Note that everywhere $(x_1/x_2)' = (x_1/x_2)(Ax)_1 - (Ax)_2 = 2x_2 - (2x_1 + 2x_3) = -4x_2 - 2$ which

shows that orbits converge to the line $Z_{1,2} = Z_{1,3}$.



Example 17. Consider $A = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 5 \\ 5 & 0 & 4 \end{pmatrix}$ determine the corresponding NE's and the phase diagrams of the replicator dynamics. Is there an ESS? $Z_{1,2}$ corresponds to $x_1 + 4x_2 = 5x_3$, $Z_{1,3}$ to $5x_2 = 4x_1 + 4x_3$ and $Z_{2,3}$ to $x_2 + x_3 = 5x_1$. These lines intersect at $E := (3/18, 8/18, 7/18)$, so this is a Nash equilibrium. Note that from the form of the indifference equations, it follows that each side of Δ is intersected by precisely two indifference lines. This, and since $\mathcal{B}R(e_i) = e_{i-1}$, implies that there just two possible positions for the Z_{ij} lines, as shown in the figure. Since $Z_{1,2}$ does not intersect $[e_1 e_2]$, the situation is as in the left figure.

The expressions for Z_{ij} were corrected on 1 Nov



Once we see this, we also obtain that orbits are rotating about the NE. Is this NE an ESS? Since the NE lies in the interior of Δ the ESS condition corresponds to $xAx < \hat{x}Ax$ for x close to the NE. Write $(x_1, x_2, x_3) = (3/8 + h_1, 8/18 + h_2, 7/18 + h_3)$. So we need to check $(x - \hat{x})Ax = (h_1 h_2 h_3)Ax < 0$. This is equivalent to $(h_1 h_2 h_3)A(h_1 h_2 h_3) < 0$. Since $h_1 + h_2 + h_3 = 0$, the last expression is equal to $h_1^2 + 5h_1h_2 + h_2^2 + 5h_2h_3 + 5h_1h_3 + 4h_3^2$. Substituting $h_3 = -h_1 - h_2$ gives that this is equal to

$$h_1^2 + 5h_1h_2 + h_2^2 + 5h_2^2 - 5h_2h_1 - 5h_2^2 - 5h_1^2 - 5h_1h_2 + 4(h_1 + h_2)^2 =$$

$$3h_1h_2 + 5h_2^2.$$

This expression does not have a constant sign for $h_1, h_2 \approx 0$. So the attracting NE is not an ESS. Nevertheless solutions converge to E. Indeed, write $x = E + h$. Then, using (8),

This calculation is incorrect. Exercise: correct it.

$$\begin{aligned} \dot{h}_1 &= (3/18 + h_1)(h_1 + 5h_2) \\ \dot{h}_2 &= (8/18 + h_2)(h_2 + 5h_3) = (8/18 + h_2)(-4h_2 - 5h_2). \end{aligned}$$

The linear part of this system is

$$\frac{1}{18}\begin{pmatrix} 3 & 15 \\ -32 & -40 \end{pmatrix}.$$

This has eigenvalues $-1.0279 \pm i \cdot 0.2341$ so the system is locally stable. To show that the system is globally stable one needs additional methods.

Exercise 1. Consider the replicator dynamics associated the following systems:

1. $A = \begin{pmatrix} 0 & 10 & 1 \\ 10 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.
2. What is the effect to the replicator dynamics $\dot{x}_i = x_i((Ax)_i - x \cdot Ax)$ of adding to the first column of A the vector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$?

1.4 Rock-paper-scissor replicator game

There is a class of systems which have only one Nash equilibrium and for which $B(e_i) = e_{i+1}$. So this suggests cyclic behaviour, and are therefore called rock-paper-scissor games. Let us consider the replicator dynamics in one example of this situation; in the general case the analysis is the same but computationally more involved.

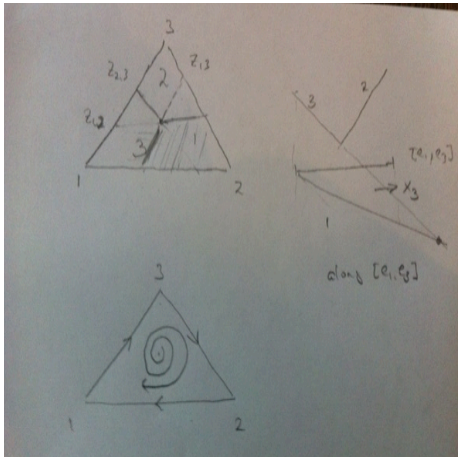
Example 18. Consider the matrix $A = \begin{pmatrix} 0 & 1 & -b \\ -b & 0 & 1 \\ 1 & -b & 0 \end{pmatrix}$ when $b > 0$. If $b = 1$ then this is a zero-sum game because then $A + A^{tr} = 0$, but otherwise it is not a zero-sum game. Show that $V := x_1x_2x_3$ is a constant of motion when $b = 1$ and draw the phase diagram. If $b \neq 1$, it is a Lyapounov function; draw the phase diagram. This game is called a **rock-paper-scissor game**. Explain why.

Ad Example 16. Note that $\mathcal{B}R(e_i) = e_{i-1}$: so the best response is cyclic. This matrix has an interior NE at $(1/3, 1/3, 1/3)$. As in Theorem 1 take $P = (x_1x_2x_3)^{1/3}$. By the calculation in that theorem, $\dot{P}/P = \hat{x} \cdot Ax - x \cdot Ax = (\hat{x} - x) \cdot Ax$. Write $x = 1/3 + h_i$. A calculation shows that $(\hat{x} - x) \cdot Ax = (b/3 - 1/3)(h_1 + h_2 + h_3) + (b - 1)(h_1h_2 + h_1h_3 + h_2h_3) = (1 - b)(h_1^2 + h_2^2 + h_1h_2)$ where in the last equality we used that $h_3 = -h_1 - h_2$. So $\dot{P}/P > 0$ when $b \in (0, 1)$ and $\dot{P}/P < 0$ when $b > 1$. So interior orbits starting at $x \neq E$, spiral out to the boundary when $b > 1$ and towards E when $b \in (0, 1)$.

Let us see whether there are other Nash equilibria. This can be done in a number of ways. One way is to calculate $\mathcal{B}R$ along $x_2 = 0$. The payoff there is $(x_2 - bx_3, -bx_1 + x_3, x_1 - bx_2) = (-bx_3, -b + (1 + b)x_3, 1 - x_3)$ where we used $x_1 = 1 - x_3$. $\mathcal{B}R(e_1) = 3$, $\mathcal{B}R(e_3) = 2$ and $x \in Z_{2,3} \cap [e_1, e_3]$ along this side implies $x_3 = (1 + b)/(2 + b)$ and then $e_2Ax = e_3Ax > 0$. Moreover, $Z_{1,3}$ holds when $x_3 = 1/(1 - b) \notin [0, 1]$. Since e_3Ax, e_1Ax are decreasing while e_2Ax increasing when going from e_1 to e_3 and also $e_1Ax = -bx_3 < 0$ along the side, it follows that the graphs of e_iAx along the side $[e_1e_3]$ as in the figure. By symmetry we obtain also the intersections of these with the other sides. It follows that there are no NE's along the sides.

Another way of concluding the positions of Z_{ij} goes as in Example 17.

Moreover, along $[e_1, e_3]$ one has $(Ax)_1 - (Ax)_3 = -bx_3 - (1 - x_3) = (1 - b)x_3 \leq 0$ as $b > 0$ and $x_3 \in [0, 1]$. So $(x_1/x_3)' < 0$ and there are no singularities along this side of the triangle. So solutions spiral out/in depending on whether $b > 1$ or $b \in (0, 1)$, and the arrows on the sides of Δ are as shown.



Lemma 5. Consider $A = \begin{pmatrix} 0 & 1 & -b \\ -b & 0 & 1 \\ 1 & -b & 0 \end{pmatrix}$ with $b > 1$. Then

$$z(T) = \frac{1}{T} \int_0^T x(t)\, dt$$

depends continuously on T and converges to some polygon with corners $A_1 = (1, b^2, b)/(1 + b + b^2)$,$A_2 = (b^2, b, 1)/(1 + b + b^2)$ and $A_3 = (b, 1, b^2)/(1 + b + b^2)$. Note that A_i, A_{i+1}, e_{i+1} are collinear. (Later on we shall see triangle is the orbit under the best response dynamics.)

Proof. Using the expression of the replicator dynamics and dividing by T gives

$$\frac{\log(x_i(T)) - \log(x_i(0))}{T} = \frac{1}{T}\sum_j a_{ij}\int_0^T x_j(t)\,dt - \frac{1}{T}\int_0^T x\cdot Ax\,dt.$$

Since $x(t)$ spends most of the time close to corners of the symplex (there the speed is small, and between corners it is large) and since $a_{ii} = 0$,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T x\cdot Ax\,dt \to 0.$$

Take a sequence $T_k \to \infty$ so that $x(T_k) \to x$ with T_k chosen to that $x_1 = 0$ and so that $x_2, x_3 > 0$ (so for this subsequence $x(T_k)$ converges to a point x somewhere in the middle of one of the sides of the triangle Δ). So $\dfrac{\log(x_i(T)) - \log(x_i(0))}{T} \to 0$ for $i = 2, 3$ and therefore

$$\sum_j a_{2j}z_j = \sum_j a_{3j}z_j = 0.$$

This means $-bz_1 + z_3 = z_1 - bz_2 = 0$. Combined with $z_1 + z_2 + z_3$ this means $z = A_1 := (b, 1, b^2)/(1 + b + b^2)$.

Permuting 1 and 2 in this argument (and 1 and 3) gives that z is equal to $A_2 = (1, b^2, b)/(1+b+b^2)$ and $A_3 = (b^2, b, 1)/(1+b+b^2)$.

Similarly, if $x(T_k)$ converges, say, to e_3 then we obtain $\dfrac{\log(x_i(T)) - \log(x_i(0))}{T} \to 0$ for $i = 3$ and therefore

$$\sum_j a_{3j}z_j = 0.$$

This means $z_1 - bz_2 = 0$. So during the long time interval when $x(T)$ stays near e_3, the average $z(T)$ travels along this segment between A_1 and A_2. □

Later on we will consider non-symmetric rock-paper-scissor games, and ask whether these can lead to chaotic dynamics.

Exercise 2. Show that A_i, A_{i+1}, e_{i+1} are collinear.

1.5 Hypercycle equation and permanence

Is it possible that orbits don't converge to the boundary and also not to a Nash equilibrium in the interior?

Let us consider an example of such a situation. Consider

$$A = \begin{pmatrix} 0 & 0 & 0 & . & . & . & k_1 \\ k_2 & 0 & 0 & . & . & . & 0 \\ 0 & k_3 & 0 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & k_n & 0 \end{pmatrix}$$

To simplify the analysis we will consider the case that $k_i = 1$. So the replicator dynamics is described by

$$\dot{x}_i = x_i(x_{i-1} - \sum_{j=1}^{n} x_j x_{j-1}). \tag{9}$$

where we cyclic notation, i.e. we take $x_0 = x_4$.

Lemma 6. This system has an interior Nash equilibrium which is stable for $n \le 4$ and unstable for $n \ge 5$.

Proof. $E = (1/n)(1, 1, \dots, 1)$ is a Nash equilibrium. The linear part of the system at this point is the matrix

$$A = \begin{pmatrix} -2/n^2 & -2/n^2 & -2/n^2 & . & . & -2/n^2 & 1/n - 2/n^2 \\ 1/n - 2n^2 & -2n^2 & -2/n^2 & . & . & . & -2/n^2 \\ -2/n^2 & 1/n - 2/n^2 & -2/n^2 & . & . & . & 2 - /n^2 \\ . & . & . & . & . & . & . \\ -2/n^2 & -2/n^2 & -2/n^2 & . & . & 1/n - 2/n^2 & -2/n^2 \end{pmatrix}.$$

This is an example of a circulant matrix

$$A = \begin{pmatrix} c_0 & c_1 & c_2 & . & . & . & c_{n-1} \\ c_{n-1} & c_0 & c_1 & . & . & . & c_{n-2} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ c_1 & c_2 & c_3 & . & . & . & c_0 \end{pmatrix}.$$

It is easy to check that the eigenvalues of such a matric are equal to

$$\gamma_k = \sum_{j=0}^{n-1} c_j \lambda^{jk}, k = 0, \dots, n-1$$

and corresponding eigenvectors

$$(1, \lambda^k, \lambda^{2k}, \dots, \lambda^{(n-1)k}), k = 0, \dots, n-1$$

where $\lambda = e^{2\pi i/n}$. In our setting this leads to $\gamma_0 = 1$ and

$$\gamma_k = \sum_{j=0}^{n-1} \frac{-2}{n^2} \lambda^{jk} + \frac{1}{n} \lambda^{(n-1)k} = \frac{\lambda^{-k}}{n}, k = 1, \dots, n-1$$

where we use that the first sum in this expression vanishes. The eigenvalue 1 with the eigenvector $(1, 1, 1, \dots, 1)$ (when $k = 0$) corresponds to the motion orthogonal to the simplex Δ so is not of interest. When $n = 3$, we get $\gamma_k = (1/2)e^{2k\pi i/3}$, $k = 1, 2$. The real parts of both these eigenvalues are negative, so the singularity is stable. When $n = 4$, we get $\gamma_k = (1/2)e^{2k\pi i/4}$, $k = 1, 2, 3$ and we see that the eigenvalues γ_1, γ_3 lie on the imaginary axis. Using a Lyapounov function, see below, we show below that in spite of this the singularity is stable when $n = 4$. For $n > 5$ there are eigenvalues with positive real part. So in this case the singularity is of saddle type, i.e. both the stable and unstable manifold of the singularity is non-empty. So the singularity is certainly not locally stable.

Let us show that the boundary of the simplex Δ is repelling. Define the Lyapounov function $P(x) = x_1, \ldots, x_{n-1}$. This function is zero on the boundary of Δ and positive in the interior of Δ. We get that

$$\frac{d}{dt}(\log P) = \sum_{i=1}^{n} \frac{\dot{x}_i}{x_i} = 1 - n \sum_{j=1}^{n} x_j x_{j-1}.$$

For $n = 2$ and $n = 3$ this function is strictly positive on the interior of Δ except at the Nash equilibrium E. When $n = 4$ then

$$\frac{d}{dt}(\log P) = \sum_{i=1}^{n} \frac{\dot{x}_i}{x_i} = 1 - n \sum_{j=1}^{n} x_j x_{j-1} =$$

$$1 - 4(x_1x_4 + x_2x_1 + x_3x_2 + x_4x_3) =$$

$$1 - 4(x_1 + x_3)(x_2 + x_4) \geq 0$$

and this is zero only when $(x_1 + x_3) = (x_2 + x_4) = 1/2$, since $(x_1 + x_3) + (x_2 + x_4) = 1$. The only invariant subset of the latter set is E and so again the Nash equilibrium is stable (and in fact attracts all orbits starting in the interior of Δ). Notice that for any arbitrary n

$$\frac{d}{dt}(\log P) = \sum_{i=1}^{n} \frac{\dot{x}_i}{x_i} = 1 - n \sum_{j=1}^{n} x_j x_{j-1} > 0$$

whenever $x \in \Delta$ is close to one of the corners e_i of Δ, and so $x(t)$ moves then away from the boundary of Δ. In the first project you are asked (based on Hofbauer and Sigmund's book) to show that orbits move away from the boundary when $n \geq 5$. So for $n \geq 5$, the attracting set is neither the boundary of Δ nor the singularity E. $\square$

1.6 Existence and the number of Nash equilibria

In this section we will show that each game has a Nash equilibrium, and in fact that for "most games" the number of Nash equilibria is odd. In fact, we will prove a result which will assign to each Nash equilibrium an index, and state that the sum of the indices is equal to $(-1)^{n-1}$ where n is the number of dimensions.

To discuss this, we will need to discuss some background on degree theory on the index of a vector field. Several results on this background will not be covered in these lectures.

To start with, let assume that M, N are smooth manifolds. These were defined in the differential equations course. But if you don't know what a manifold is, then think of for example $M = \mathbb{R}^n$, M is an open ball in $\mathbb{R}^n$, $M = S^1$, $M = S^2$ or $M = T^2$.

Moreover, let $f\colon M \to N$ be a smooth map. We say that y is a *regular value* if $f^{-1}(y) \neq \emptyset$ and if for each $x \in f^{-1}(y)$ the map f locally smoothly invertible near x, i.e. Df_x is an invertible matrix. In this case, define

$$\operatorname{sign} Df_x = \begin{cases} +1 & \text{if it is locally orientation preserving} \\ -1 & \text{if it preserves orientation.} \end{cases}$$

Definition. Let $f\colon M \to N$ be a smooth map and that y is a regular value. Then the degree of f at a point $y \in f(M) \setminus f(\partial M)$ is equal to

$$\deg(f; y) = \sum_{x \in f^{-1}(y)} \operatorname{sign} Df_x.$$

Example 19. Let $S^1 = \mathbb{R}/\mathbb{Z}$ and $f\colon S^1 \to S^1$ be defined by $x \mapsto 2 = nx$. Then $\deg(f, y) = n$ for each y. Let M be an open ball in $N := \mathbb{R}^n$ and define $f(x) = -x$. Then $\deg(f, y) = (-1)^n$ for each $y \in N = f(M)$.

Theorem 2. The degree of a map has the following useful properties:

- The degree $\deg(f, y)$ of $f\colon M \to N$ is the same for each regular value y of f, see figure in lecture. So this why we speak also of $\deg(f)$.
- If $f_t\colon M \to N$ is a family of smooth maps depending smoothly on t, then $\deg(f_0) = \deg(f_1)$.

Definition. Consider $X\colon \mathbb{R}^n \to \mathbb{R}^n$, and assume x_0 is an isolated zero of X. We will view X as a vector field, so at each point $x \in \mathbb{R}^n$ we have a vector $X(x)$. Take a small sphere S centered at x_0 on which X has no zeros, and define the map

$$f\colon S \to S^{n-1}, \text{ by } f(x) = \frac{X(x)}{|X(x)|}.$$

Then the degree of X at x_0 is defined as

$$\text{ind}(X, x_0) := \deg(f).$$

The same definition applies if X is a vector field on a manifold.

Example 20. The index of a saddle point in $\mathbb{R}^2$ is -1, and of a source and a sink is 1.

Lemma 7. Assume that X is a vector field, and x_0 an isolated singularity and that its linearisation $A := DX(x_0)$ is non-singular (i.e. invertible). Then $\text{ind}(X, x_0)$ is equal to the sign of the determinant of A.

In particular, we have $\text{ind}(-X, x_0) = (-1)^n \cdot \text{ind}(X, x_0)$ where n is the dimension.

It is easy to check this in dimension two or for linear vector fields. The general case can be seen by deforming the vector field continuously to the linear one, without introducing new singularities.

Example 21. Consider the vector field $X(x) = -x$ on $\mathbb{R}^n$. This corresponds to the differential equation $x_i' = -x_i$, $i = 1, \ldots, n$. Then according to the previous theorem, $\text{ind}(X, 0) = (-1)^n$. Moreover, the associated map is equal to $f(x) = -x$, and so again $\text{deg}(f, 0) = (-1)^n$.

The following remarkable theorem is related to the famous Brouwer fixed point theorem.

Theorem 3 (Poincaré-Hopf theorem)**.** Let X be a vector field which is defined on a compact manifold M (you may assume that M is a compact subset of $\mathbb{R}^n$), and assume that if M has a non-empty boundary then for each $x \in \partial M$ one has that $X(x)$ points outwards.

Then

$$\sum_{x, X(x)=0} i_X(x) = \chi(M)$$

where $\xi(M)$ is the Euler characteristic of M.

In this course we won't develop the machinery to compute (or even to formally define) the Euler characteristic of a space. For this you need some homology theory, a subject which is covered in most courses on algebraic homology, and so outside the scope of this course. However, let us give some examples.

Example 22. The sphere S^2 in $\mathbb{R}^3$ has Euler characteristic 2. A surface which is made up of a sphere with g handles, has Euler characteristic $2 - 2g$. So for example the torus has Euler characteristic 0 and the pretzel Euler characteristic -2. In fact, assume that you describe a surface as a convex polyhedron. Then its Euler characteristic $\xi = V - E + F$ where V, E, F are the number of vertices, edges and faces. For example, for a cube $V = 8, E = 12, F = 6$ and so $\xi = 2$ while for a tetrahedron, $V = 4, E = 6, F = 4$ and so again $\xi = 2$.

Similarly, an open or closed ball B in $\mathbb{R}^n$ has Euler characteristic 1 whereas the sphere S^n in $\mathbb{R}^{n+1}$ has Euler characteristic $1 + (-1)^n$.

Example 23. The above theorem implies the *Brouwer's fixed point theorem* if we assume that the map involved is smooth. This theorem says that any continuous map $f\colon B \to B$ from a ball in $\mathbb{R}^n$ has a fixed point. Let us assume that f is smooth, B is the unit ball and by contradiction assume that f has no fixed point. Then we can define the vector field $X(x)$ defined by $X(x) = x - f(x)$ has no zeros and points along the boundary to the exterior of B. But this contradicts the Poincaré-Hopf theorem as $\xi(B) = 1$.

Example 24. The above theorem also implies the so-called hairy ball theorem: If X is a vector field on S^2 then

$$\sum_{x, X(x)=0} i_X(x) = 2.$$

In particular X has at least one zero. The reason this is called the hairy ball theorem is that it implies that a hairy ball has to have places where the "hair sticks up". Note that the above theorem also implies that it is impossible to have a vector field on S^2 with just one saddle point.

Application to game theory

We say that a singularity x_0 of a vector field X is *regular* if the linear part $A = DX(x_0)$ is invertible, and say that a game is *regular* if at each Nash equilibrium $\bar{x}$, the replicator dynamics has a regular singularity (i.e. the linearisation is invertible - so no zero eigenvalue).

Remark 1. Assume that x_0 is a regular singularity of the vector field X and let X_λ is a family of vector fields depending differentiably on λ with $X_0 = X$. Then by the implicit function theorem, there exists a function $\lambda \to x_0(\lambda)$ so that $X(x_0(\lambda)) = 0$. (So the singularity moves smoothly as the parameter varies.)

Theorem 4. Each $n \times n$ matrix A has at least one Nash equilibrium. Moreover,

1. if A is a regular game, then the number of its Nash equilibria is odd.

2. Consider a Nash equilibrium $\bar{x}$ of the replicator dynamics $\dot{x} = X(x)$ on the boundary of Δ and let $B = DX(\bar{x})$ its linear part. Then any eigenvalue corresponding to any eigenvector of B which is transversal to the boundary of Δ is negative. Hence the stable manifold of $\bar{x}$ points into the interior of Δ, and the unstable manifold of $\bar{x}$ is either empty or fully contained in $\partial\Delta$.

3. Most $n \times n$ matrices are regular.

Proof. Consider the following slight modification of a replicator equation:

$$\dot{x}_i = x_i((Ax)_i - x \cdot Ax - n\epsilon) + \epsilon. \tag{10}$$

and let X_ϵ be the vector field defined by the r.h.s. of this expression. Along $\partial\Delta$, the vector field $X_\epsilon(x)$ has no singularities, and points into the simplex Δ. This means that along $\partial\Delta$ the vector field $-X_\epsilon(x)$ points outwards. So, by the Poincaré-Hopf theorem, the sum of the indices of the singularities of $-X_\epsilon$ is equal to 1. Now note that X and $-X$ have the same singularities, and by Lemma 7 at each singularity x_0 we have $\text{ind}(-X_\epsilon, x_0) = (-1)^{n-1}\text{ind}(X_\epsilon, x_0)$ because the dimension of Δ is $n-1$. It follows that for each $\epsilon > 0$, the sum of the indices of the singularities of (10) is equal to $(-1)^{n-1}$.

For any singularity $p(\epsilon)$ of (10) we have

$$(Ap(\epsilon))_i - p(\epsilon) \cdot Ap(\epsilon)) = n\epsilon - \frac{\epsilon}{p_i(\epsilon)}.$$

Hence, for any limit point $\bar{p}$ of $\lim_{\epsilon \to 0} p(\epsilon)$ we have that

$$(A\bar{p})_i \leq \bar{p} \cdot A\bar{p}$$

and so $\bar{p}$ is a Nash equilibrium.

Moreover, if all singularities of the replicator system X_0 are regular, then X_0 has finitely many singularities (each one is isolated). Each of these singularities moves smoothly with ϵ and remains a singularity of X_ϵ, i.e. of (10). Moreover, the number singularities of (10) remains the same for all $\epsilon \geq 0$ small. This proves the first assertion of the lemma.

To prove the 2nd assertion, let us consider a singularity $\bar{x}$ on the boundary, i.e. with $\bar{x}_i = 0$. Because of the form of the equation, the i-row of the linearisation B is of the form $(0\,0\,\ldots z_i \ldots 0\,0)$ where $z_i = (A\bar{x})_i - \bar{x} \cdot A\bar{x}$ appears on the i-position and the other terms are zero. It follows that any eigenvector with a non-zero i-component has eigenvalue z_i. Since we assumed that the system is regular and it is a Nash equilibrium, we have $z_i < 0$. Hence the 2nd claim holds.

It is not so hard to prove the 3rd assertion, but we will not do this here. □

Example 25. In Example 15 we had three Nash equilbria: e_1, E and $[e_1, e_3] \cap Z_{1,3}$. The first one is a sink, the 2nd a source, and the final one a saddle, so with index $+1, +1, -1$. The sum of these numbers is equal to $+1 + 1 - 1 = 1 = (-1)^{3-1}$. In several other examples we had a unique NE which was a sink or source in the interior (or a centre) and so there the theorem also holds.

Exercise 3. Give a heuristic argument which shows that if a game has only regular singularities, then the Nash equilibria on the boundary move into the interior of Δ and the other singularities move out of Δ.

2 Iterated prisoner dilemma (IRP) and the role of reciprocity

In this chapter we will consider the prisoner dilemma game and donation game from the introduction of these notes.

In the latter game the payoff is

$$\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}$$

where we assume $b > c > 0$. (This game works as follows: If a player pays c into the scheme the other player receives a benefit of b.) So if you cooperate and the other player too then you receive $b-c$, but if you do but the other person not, then you loose $-c$.

Of course this is a special case of the prisoner dilemma game

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} \text{ with } T > R > P > S.$$

Note that the 2nd strategy dominates the first one (i.e. the 2nd row dominates the first one).

Of course if this game is repeated exactly 100 times, then by induction you can deduce that best strategy is both players is to never donate (i.e. always defect).

In this chapter we will consider various more realistic scenarios in which you don't know how many times this game is repeated.

Various alternative strategies, i.e. Tit for Tat strategies are then discussed.

2.1 Repeated games with unknown time length

Let us assume that after each round there is a probability w that the game is repeated at least one more round, where $w \in [0,1]$.

So the probability of the game taking exactly n rounds is $w^n(1-w)$. This means that the expected duration of the game is

$$1(1-w) + 2w(1-w) + \ldots nw^{n-1}(1-w) + \cdots = \frac{1}{1-w}.$$

Let us assume that the payoff at time n is equal to A_n. Then the expected value of the total payoff is

$$\sum_{n=0}^{\infty} w^n(1-w)[A_0 + \cdots + A_n].$$

It is easy to see that this is equal to the convergent series

$$A(w) := A_0 + wA_1 + w^2A_2 + \ldots.$$

Since A_n is finite, this sum exists. As the expected duration of the game is $1/(1-w)$, the average payoff per round is therefore

$$(1-w)A(w).$$

In the limiting case $w = 1$, the above sum in the definition of $A(w)$ does not converge, but instead we can look at the limit of the average payoff:

$$\frac{A(0) + \cdots + A(n)}{n+1}.$$

Let us the case where the players consider three strategies: AllC, AllD, TFT. This means Always Cooperate, Always Defect or Tit For Tat (TFT means cooperate if and only if the other player cooperated last time).

Let us assume that in the TFT strategy, in the first round the player cooperates. Then the payoff matrix is

$$\frac{1}{1-w}\begin{pmatrix} b-c & -c & b-c \\ b & 0 & b(1-w) \\ b-c & -c(1-w) & b-c \end{pmatrix}$$

where strategies are AllC, AllD, TFT.

To see this, consider the situation where both players cooperate. Then the payoff $A_n = b - c$ and so $A(w) = (b - c)/(1 - w)$. If both players play TFT then they will keep cooperating, and so the payoff is again $A(w) = (b - c)/(1 - w)$. If I play TFT and the other player plays AllD, then $A_0 = -c$ and $A_n = 0$, so $A(w) = -c$. On the other hand, if I play AllD and the other player TFT, then $A_0 = b$ and from then on $A_n = 0$, so $A(w) = b$.

When we let $w \to 1$ then we get the case where the game is repeated infinitely often.

Instead of the above situation, let us change the payoff matrix, by adding to each column a multiple of the vector $\mathbb{1}$. As we have seen this does not change the replicator dynamics. Let's do this so the 2nd row consists of all 1's, and then multiply the matrix by $(1 - w)$. This gives

$$\hat{A} = \begin{pmatrix} -c & -c & bw - c \\ 0 & 0 & 0 \\ -c & -c(1-w) & bw - c \end{pmatrix}.$$

Then

$$(Ax)_1 = -c + wbx_3\,,\ (Ax)_2 = 0 \text{ and } (Ax)_3 = (Ax)_1 + wcx_2.$$

Note that the best response is always e_2 if $w < c/b$ but that if $w > c/b$ then $\mathcal{B}R(e_1) = e_2$, $\mathcal{B}R(e_2) = e_2$ and $\mathcal{B}R(e_3)$ is multivalued an equal to $\mathcal{B}R(e_3) =< e_1, e_3 >$. So let us assume that $w > c/b$. Note that the 3rd row is dominating the 1st one when $x_2 > 0$ and that $(A\tilde{x})_1 = (A\tilde{x})_2 = (A\tilde{x})_3$ holds for $\tilde{x}$ with $\tilde{x}_3 = c/wb$ and $\tilde{x}_2 = 0$. Using slightly annoying calculations we get

$$x{\cdot}Ax = (Ax)_3 - x_2 g(x_3), \text{ where } g(x_3) = w(b-c)x_3 - c(1-w).$$

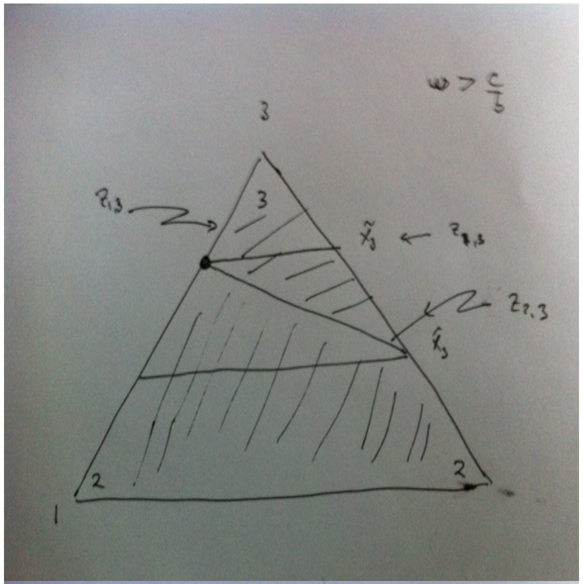
Hence $\dot{x}_3$ along the line $g(x_3) = 0$ and so this lines is invariant. Note that

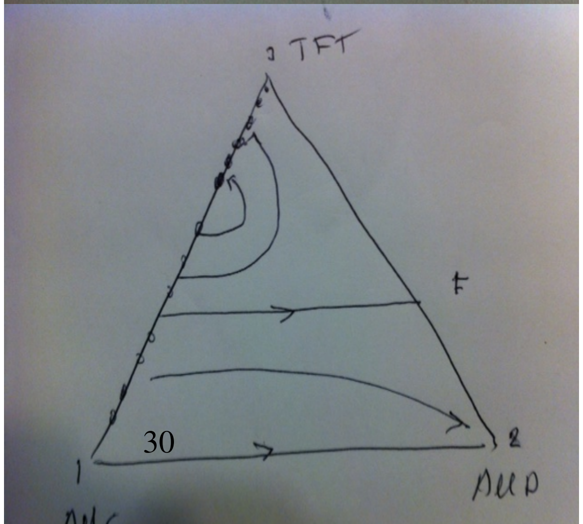
$$g(\hat{x}_3) = 0 \iff \hat{x}_3 = \frac{(1-w)c}{w(b-c)}$$

which is < 1 iff $w > c/b$. Note that $\hat{x}_3 < \tilde{x}_3$ whenever $w > c/b$ because then $(1 - w)/(b - c) < 1/b$.

Along $x_2 = 0$, we have $\dot{x}_2 = 0$ and $(Ax)_3 - x \cdot Ax = x_2 g(x_3) = 0$, so $\dot{x}_3 = \dot{x}_1 = 0$. So $x_2 = 0$ consists of singularities. The singularities with $x_2 = 0$ and $x_3 \geq c/wb$ are attracting (and Nash equilibria) and the the singularities with $x_2 = 0$ and $x_3 < c/wb$ are not Nash equilibria.

There are no interior singularities because $(Ax)_3 > (Ax)_1$ when $x_2 > 0$.





2.2 Random versions of AllC, AllD and TFT

Let us consider a modification of the previous set-up, in which a player makes a probabilistic response to the other players position.

More precisely, in each round there are four possibilities: (C, C), (C, D), (D, C), (D, D). Let us label these as 1, 2, 3 and 4 with payoff for the first player of R, S, T, P, and let $x(n)$ be the probability that each of these is played at time n (so this is a probability vector in $\mathbb{R}^4$).

Let us describe the probabilistic set-up by vectors (f, p, q) and (f', p', q') for each player. For example, a player that chooses the TFT strategy from the previous section is described by $(f, p, q) = (1, 1, 0)$, which means that he plays C in the first round ($f = 1$), will definitely reciprocate a C with a C ($p = 1$ but punish a D with a D ($q = 0$).

More formally, $f, f' \in [0, 1]$ gives the probability that the first and 2nd player Similarly, let p, q be the probability of player I responding in the next round with C when player II plays respectively C, D. So assume that player I cooperates with probability $c(n)$ in round n, then the probability of player II cooperating in round $n + 1$ is equal to

$$c'(n+1) = p'c(n) + q'(1 - c(n)) = q' + \rho' c(n)$$

where $\rho' = p' - q'$. The probability of player I cooperating in round $n + 2$ is equal to

$$c(n+2) = q + \rho c'(n+1) = A + uc(n)$$

where $\rho = p - q$, $A = q + \rho q'$ and $u = \rho\rho'$. It follows that

$$c(2n) = v + u^n(f - v)$$

where f is the probability of player I choosing C in the first round and $v = \dfrac{q + \rho q'}{1 - \rho\rho'}$. A similar equation holds for $c(2n +$

1). In the special case of the donation game (with coefficients b, c), and again considering the situation of a probability $1 - w$ after each round of terminating the game, we obtain the average payoff per round of strategy (f, p, q) against (f', p', q') is

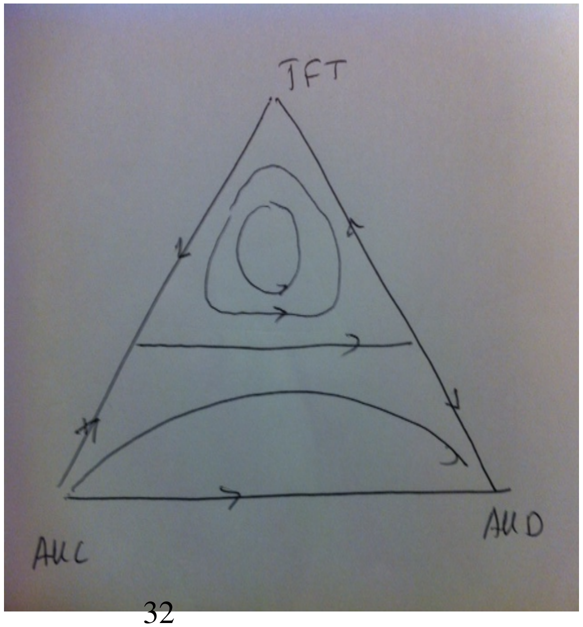
$$\frac{-c(e + w\rho e') + b(e' + w\rho' e)}{1 - uw^2}$$

where $e = (1 - w)f + wq$, $e' = (1 - w)f' + wq'$.

Let us consider the three possible strategies: $e_1 = (1 - \epsilon, 1 - \epsilon, 1 - \epsilon)$, $e_2 = (k\epsilon, k\epsilon, k\epsilon)$ and $e_3 = (1 - \epsilon, 1 - \epsilon, k\epsilon)$. Using some simplifications we obtain the (normalised) payoff matrix

$$A = \begin{pmatrix} 0 & -1 & \delta\sigma \\ 1 & 0 & -\kappa\sigma \\ \delta & -\kappa & 0 \end{pmatrix}$$

where $\delta = w\epsilon$, $\kappa = 1 - w + wk\epsilon$, $\sigma = \dfrac{b\theta - c}{c - c\theta}$ and $\theta = w(1-(k+1)\epsilon)$. This corresponding to the following replicator dynamics:



3 Other game dynamics

Let us consider several dynamical systems which are related to the replicator dynamics. As a test case we shall consider the rock-scissor-paper pay-off matrix

$$A = \begin{pmatrix} 0 & -b & a \\ a & 0 & -b \\ -b & a & 0 \end{pmatrix} \tag{11}$$

Remember that this system as a unique Nash equilibrium at $E = (1/3)\mathbb{1}$. Moreover, writing $x = E + z$ then $\sum_i z_i = 0$ we have

$$z \cdot Az = (a - b)(z_1 z_1 + z_2 z_3 + z_1 z_2) = \frac{(b - a)}{2}[z_1^2 + z_2^2 + z_3^2]$$

This means that E is an ESS when $0 < b < a$ and it is NOT an ESS when $0 < a < b$. Note that orbits move away/towards the Nash equilibrium if $z \cdot Az < 0$ resp. $z \cdot Az > 0$ for all $z \in \mathbb{R}^3$ with $\sum_i z_i = 0$. (This follows as in Theorem 1.)

3.1 Best response dynamics

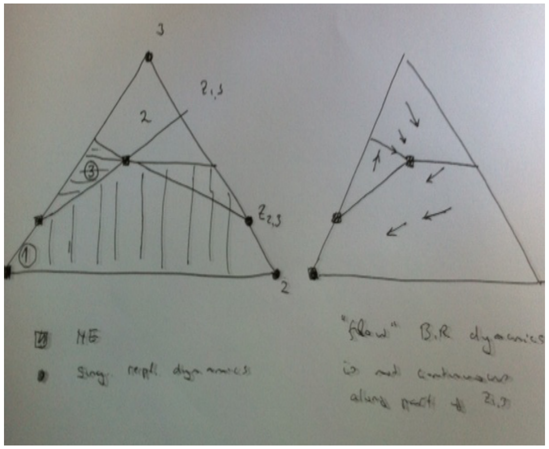
Define as before $\mathcal{B}R(x) = \arg\max_y yAx$ and let

$$\dot{x} = \mathcal{B}R(x) - x. \tag{12}$$

This is called the best response dynamics, and this is much older than replicator dynamics. Note that $\mathcal{B}R(x)$ is a non-empty convex set. In fact, it is upper semi-continuous. (What does that mean?) It follows from general principles that the the best response dynamics has a solution in the sense that there exists a solution $t \mapsto x(t)$ of (12) with $x(0) = x_0$ so that $t \mapsto x(t)$ is almost every where differentiable.

Example 26. Consider Example 11 from above and take $V(x) = \max_i Ax$. Then $V(x) = e_i \cdot Ax$ where $i = i(x) = \mathcal{B}R(x)$ is piecewise constant. So $\dot{V} = e_i \cdot A \cdot x = e_i \cdot A(e_i - x) = -e_i Ax = -V$ except at the Nash equilibrium E. Note that $V(E) = (a-b)/3$. So when $a > b$ then $V(E) > 0$ and $V(x) \geq V(E)$ for all $x \in \Delta$. Indeed, write $x = z + E$ with $\sum z_i = 0$; therefore $\sum_i (Az)_i = 0$ and $\max(Az)_i \geq 0$. It follows that orbits reach E in finite time. If $a < b$ then $V(E) < 0$ and so the solution does NOT converge to Nash, but to the set where $V = 0$, which is a triangle, called the Shapley triangle.

Example 27. Let us consider the $\mathcal{B}R$-dynamics from the remaining example 15 in Section 1.3, where $A = \begin{pmatrix} 0 & 6 & -4 \\ -3 & 0 & 5 \\ -1 & 3 & 0 \end{pmatrix}$. Notice that the $\mathcal{B}R$-dynamics is multivalued along the indifference lines. Along the segment of the line $Z_{2,3}$ where the regions 2 and 3 meet in example 15, there is a unique continuous extension of the flow, see the lecture. Along the segment of the line $Z_{1,3}$ where regions 1 and 3 meet the 'flow' is non-continuous. See lectures.



3.2 Logit dynamics

Although the best response dynamics has major benefits, since the orbits are piecewise straight arcs, the corresponding solutions do not depend continuously on initial conditions and one does not have the flow property. For this reason some people prefer to use a smoother version of the best response dynamics. One possible definition goes as follows. Consider the 'logit' function

$$L\colon \mathbb{R}^n \to \Delta \text{ defined by } L_k(u) = \frac{e^{u_k}}{\sum_j e^{u_j}}, k = 1, \ldots, n.$$

Corresponding to this one has the logit dynamics

$$\dot{x} = L(Ax/\epsilon) - x.$$

3.2.1 First motivation for logit dynamics

Note that when $\epsilon \to 0$ and the k-th component of u is larger than its other components then $L(u/\epsilon) \to e_k$. So when $\epsilon > 0$ is small, then one can think of $L(Ax/\epsilon)$ as an approximation of the $\mathcal{B}R(x)$. Note that $L(Ax/\epsilon)$ is a unique probability vector whereas $\mathcal{B}R(x)$ can be set-valued.

3.2.2 Second motivation for logit dynamics

Another way to obtain (or motivate) the logit dynamics goes as follows. Let us consider random variables $\epsilon_i\colon \mathbb{R} \to \mathbb{R}$ and the function

$$\hat{L}\colon \mathbb{R}^n \to \Delta \text{ defined by } \hat{L}_k(u) = Prob(u_k + \epsilon_k \geq u_j + \epsilon_j \forall j).$$

Let us show that if we choose ϵ_i appropriately that $\hat{L}$ is equal to L. Define $F(x) = \exp(-\exp(-x))$ is monotone increasing and converges to 0 and 1 as $x \to -\infty$ resp. $x \to \infty$. So F is a cumulative distribution function and to say that $\epsilon \sim F$

A real valued random variable ϵ is simply a measurable function from a probability space Ω to $\mathbb{R}$. The space Ω has associated to it a collection of subsets $\mathcal{F}$ which are called the measurable sets, and a measure $\mu\colon \mathcal{F} \to [0, 1]$ with the main property that $\mu(\cup A_i) = \sum_i \mu(A_i)$ when $A_i \in \mathcal{F}$ are a *countable* and mutually disjoint.

means that $Prob(a \leq \epsilon \leq b) = F(b) - F(a)\,dx$. (Or in other words, $Prob(a \leq \epsilon \leq b) = \int_a^b f(x)\,dx$ where f is the derivative of F. This is called the probability distribution function which in this case is $e^{-x}e^{-e^{-x}}$.) Note that, by rewriting and considering ϵ_k for the moment as given, and using that the ϵ_j are independent, we get

$$\begin{aligned} \hat{L}_k(u)|\epsilon_k &= Prob(\epsilon_j \leq u_k - u_j + \epsilon_k \; \forall j \neq k) \\ &= \prod_{j \neq k} \exp(-\exp(u_j - u_j + \epsilon_k)). \end{aligned}$$

Now in fact ϵ_k is not given, but has probability distribution $e^{-s}e^{-\exp(-s)}$ and so substituting $s = \epsilon_k$ and then $t = e^{-s}$,

$$\begin{aligned} \hat{L}_k(u) &= \int_{-\infty}^{\infty} \prod_{j \neq k} e^{-\exp(u_k - u_j + s)} e^{-s} e^{-e^{-s}}\,ds. \\ &= \int_{-\infty}^{\infty} e^{-e^{-s}\sum_j \exp(u_k - u_j)} e^{-s}\,ds \\ &= \int_0^{\infty} e^{-t(\sum_j \exp(u_k - u_j))}\,dt \\ &= \left. \frac{e^{-t\sum_j \exp(u_k - u_j)}}{-\sum_j \exp(-(u_k - u_j))} \right|_{t=0}^{\infty} \\ &= \frac{e^{u_k}}{\sum_j e^{u_j}}. \end{aligned}$$

This gives the claimed result.

Note that singularities of the logit dynamics are not necessarily Nash equilibria (and vice versa).

Example 28. Let us for example consider A from (11) as above. Then AE is $(a-b)\mathbb{1}$ and $L(AE/\epsilon) = \dfrac{e^{(a-b)/\epsilon}}{\sum_j e^{(a-b)/\epsilon}} = E$. So in this case the Nash equilibrium is indeed a singularity of the logit dynamics. The linearisation of the Nash equilibrium can be done explicitly.

4 Two player games

So far we looked at a game with one population with different strains, and analysed whether a particular make-up $\bar{x}$ is 'optimal', in the sense that is a Nash or an ESS equilibrium.

A more general situation is when there are two populations which are competing. In this case we have two matrices A, B and assume that the positions of the two populations are determined by two probability vectors x and y, the **payoff** and best-response maps for the two populations is respectively is

$$\begin{array}{ll} x \cdot Ay\,, & \mathcal{B}R_A(y) = \arg\max_{x\in\Delta} x \cdot Ay, \\ y \cdot Bx\,, & \mathcal{B}R_B(x) = \arg\max_{y\in\Delta} y \cdot Bx. \end{array} \tag{13}$$

We then say that $(\hat{x}, \hat{y})$ is a *Nash equilibrium* iff

$$\hat{x} \in \mathcal{B}R_A(\hat{y}) \text{ and } \hat{y} \in \mathcal{B}R_B(\hat{x}).$$

Note that it is not necessary that A, B are square matrices (and that the dimension of the probability spaces for x and y are equal). So A would be a $n \times m$ matrix and B a $m \times n$ matrix, which means that player A has n strategies and B has m strategies to choose from.

Note that if $A = B^{tr}$ then $y \cdot Bx = x \cdot Ay$ and so $\hat{x} \in \mathcal{B}R_A(\hat{x})$ implies also $\hat{x} \in \mathcal{B}R_B(\hat{x})$. So $(\hat{x}, \hat{x})$ is a NE for the game determined by (A, B). This is the *symmetric case* which we considered so far. It follows from Theorem 4 that any symmetric game (one with $A = B^{tr}$) has a symmetric Nash equilibrium. If $A + B^{tr} = 0$ then we say that the game is *zero-sum*.

In fact, there is also **2nd convention** for defining the payoff of two players, namely to define the payoff and best response functions, namely

$$\begin{array}{ll} x \cdot Ay\,, & \mathcal{B}R_A(y) = \arg\max_{x\in\Delta} x \cdot Ay \quad \text{and} \\ x \cdot By\,, & \mathcal{B}R_B(x) = \arg\max_{y\in\Delta} x \cdot By. \end{array} \tag{14}$$

With *this convention* a zero-sum game corresponds to $A+B=0$. The benefit of the 2nd convention becomes clear in the following example:

Example 29. Let us consider the situation where both players have two strategies and so the payoff matrices are 2×2: $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$ and $B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$. If we use the convention from (14) then one can combine these two matrices using the following notation $\begin{pmatrix} (a_1,b_1) & (a_2,b_2) \\ (a_3,b_3) & (a_4,b_4) \end{pmatrix}$. This corresponds to

Payoff's	Player B chooses left	Player B chooses right
Player A chooses top	(a_1,b_1)	(a_2,b_2)
Player A chooses bottom	(a_3,b_3)	(a_4,b_4)

,

4.1 Two player replicator dynamics

The replicator dynamics corresponding to two populations is defined as

$$\begin{aligned} \dot{x}_i &= x_i((Ay)_i - x \cdot Ay) \\ \dot{y}_j &= y_j((Bx)_j - y \cdot Bx) \end{aligned}$$

if we use the first convention for A, B as in 13.

Let us consider the 2x2 case, with payoff matrices $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ and use the first convention for the moment. This gives

$$\begin{aligned} \dot{p}_1 &= p_1[a_{11}q_1 + a_{12}q_2 - p_1(a_{11}q_1 + a_{12}q_2) - p_2(a_{21}q_1 + a_{22}q_2)] \\ &= p_1(1-p_1)[\alpha_1 - q_1(\alpha_1 + \alpha_2)] \\ \dot{q}_1 &= q_1[b_{11}p_1 + b_{12}p_2 - q_1(b_{11}p_1 + b_{12}p_2) - q_2(b_{21}p_1 + b_{22}p_2)] \\ &= q_1(1-q_1)[\beta_1 - p_1(\beta_1 + \beta_2)] \end{aligned}$$

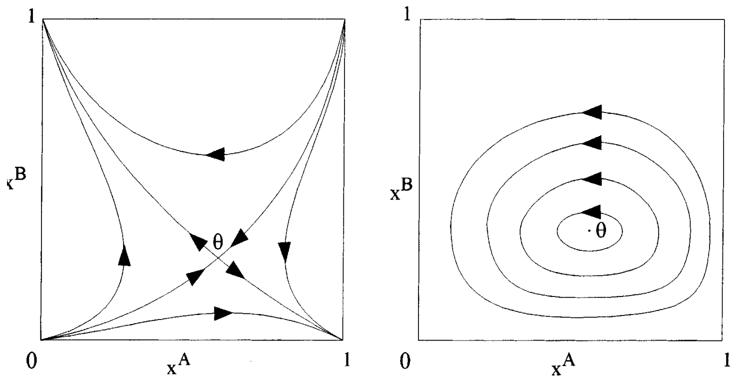
where

$$\begin{aligned} \alpha_1 = a_{12} - a_{22}, &\quad \alpha_2 = a_{21} - a_{11} \\ \beta_1 = b_{12} - b_{22}, &\quad \beta_2 = b_{21} - b_{11}. \end{aligned}$$

It turns out that there are three possibilities:

Proposition 1. There are three possibilities for a 2×2 replicator dynamics system (apart from the degenerate case), namely
(i) $\alpha_1\alpha_2 > 0, \beta_1\beta_2 > 0, \alpha_1\beta_1 > 0$ (battle of the sexes),
(ii) $\alpha_1\alpha_2 < 0$ or $\beta_1\beta_2 < 0$ (dominated strategy),
(iii) $\alpha_1\alpha_2 > 0, \beta_1\beta_2 > 0, \alpha_1\beta_1 < 0$ (zero sum case). The dynamics in the first and last one is as drawn below.

Include prove??



Exercise 4. Give examples which correspond to these figures.

4.2 A 3×3 replicator dynamics systems with chaos

A well-known example of a two-player game is

$$A = \begin{pmatrix} \epsilon_x & -1 & 1 \\ 1 & \epsilon_x & -1 \\ -1 & 1 & \epsilon_x \end{pmatrix} B = \begin{pmatrix} \epsilon_y & -1 & 1 \\ 1 & \epsilon_y & -1 \\ -1 & 1 & \epsilon_y \end{pmatrix}$$

Note that this game is zero-sum if $A + B^{tr} = 0$ so when $\epsilon_x + \epsilon_y = 0$. For this game numerical investigations by Sato, Akiyama and coworkers show chaos.

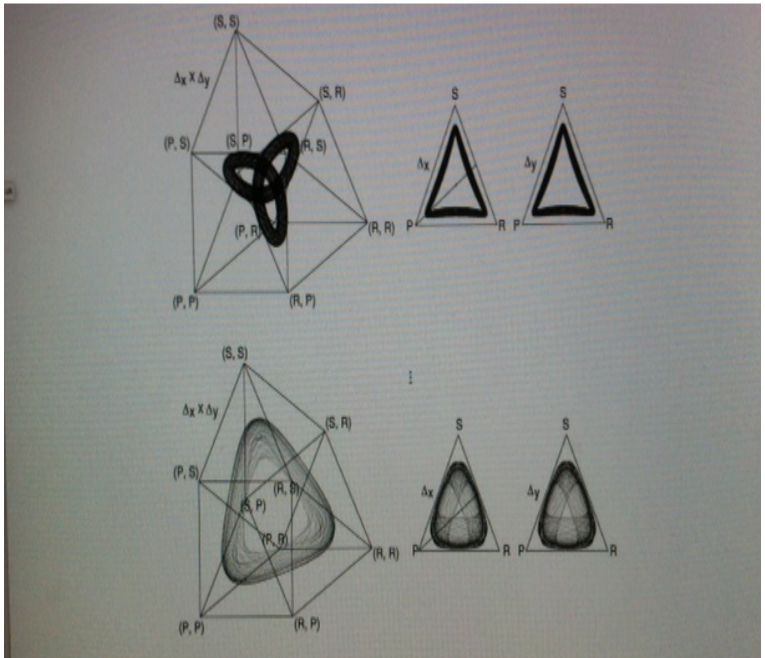
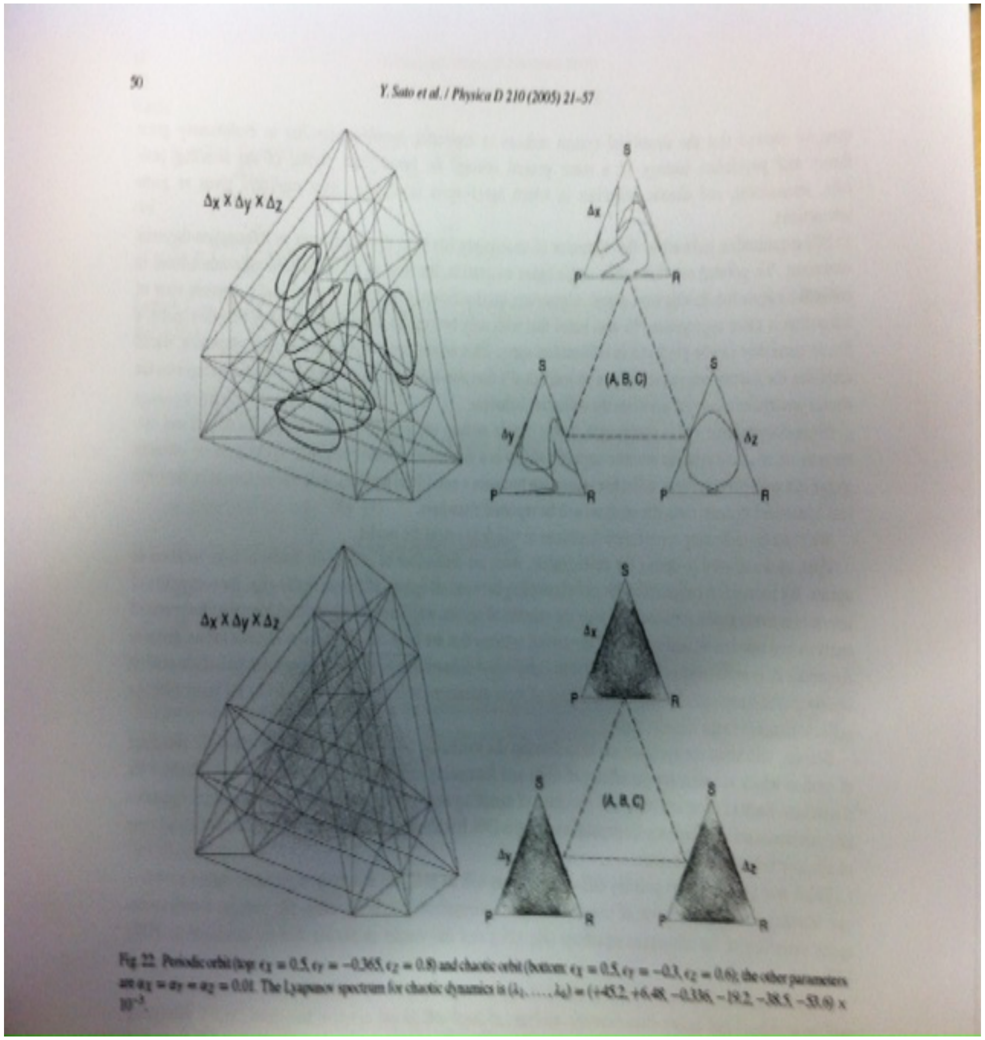


Fig. 12. Quasiperiodic tori: collective dynamics in Δ (left column) and individual dynamics projected onto Δ_X and Δ_Y, respectively (right two columns). Here $\epsilon_X = -\epsilon_Y = 0.0$ and $\alpha_X = \alpha_Y = 0$. The initial condition is (A): $(\mathbf{x}, \mathbf{y}) = (0.26, 0.113333, 0.626667, 0.165, 0.772549, 0.062451)$ for the top and (B): $(\mathbf{x}, \mathbf{y}) = (0.05, 0.35, 0.6, 0.1, 0.2, 0.7)$ for the bottom. The constant of motion (Hamiltonian) is $E = 0.74446808 \equiv E_0$. The Poincaré section used for Fig. 14 is given by $x_1 = x_2$ and $y_1 < y_2$ and is indicated here as the straight diagonal line in agent X's simplex Δ_X.



50 Y. Sato et al. / Physica D 210 (2005) 21–57

Fig. 22. Periodic orbit (top: $\epsilon_X = 0.5, \epsilon_Y = -0.365, \epsilon_Z = 0.8$) and chaotic orbit (bottom: $\epsilon_X = 0.5, \epsilon_Y = -0.3, \epsilon_Z = 0.6$); the other parameters are $\alpha_X = \alpha_Y = \alpha_Z = 0.01$. The Lyapunov spectrum for chaotic dynamics is $(\lambda_1, \ldots, \lambda_6) = (+45.2, +6.48, -0.336, -19.2, -38.5, -53.6) \times 10^{-3}$.

4.3 Two player best response dynamics

The best response dynamics corresponding to two populations is

$$\begin{aligned} \dot{x} &= \mathcal{B}R_A(y) - x \\ \dot{y} &= \mathcal{B}R_B(x) - y \end{aligned}$$

Example 30. Let us consider the example of $\begin{pmatrix} (-1,1) & (0,0) \\ (0,0) & (-1,1) \end{pmatrix}$ (so we use convention 14). Here both players have opposite interests (the sum of the payoff's is always zero) and there is a unique interior NE, namely at $E = (1/2, 1/2) \times (1/2, 1/2)$. Let us show that in this case solutions go to this NE. Take

$$V(x,y) = \mathcal{B}R_A(y)Ay + xB\mathcal{B}R_B(x).$$

Notice

$$\mathcal{B}R_A(y)Ay \geq xAy \text{ and } xB\mathcal{B}R_B(x) \geq xBy = -xAy.$$

It follows that $V(x,y) \geq 0$. Moreover, at $E = (E^A, E^B)$ we have $V(E) = \mathcal{B}R_A(E^B)AE^B + E^A B\mathcal{B}R_B(E^A) = E^A A E^B + E^A B E^B = 0$. Moreover,

$$\begin{aligned} \dot{V} &= \mathcal{B}R_A(y)A\dot{x} + \dot{x}B\,\mathcal{B}R_B(x) \\ &= \mathcal{B}R_A(y)A(\mathcal{B}R_B(x) - y) + (\mathcal{B}R_A(y) - x)B\,\mathcal{B}R_B(x) \\ &= -V \end{aligned}$$

where in the last step we used $A + B = 0$. It follows that $V(x(t), y(t)) = e^{-t}V(x(0), y(0))$. This means that orbits tend exponentially fast to the Nash equilibrium E. The orbits spiral to the NE as in the 2nd figure in the picture below.

4.4 Convergence and non-convergence to Nash equilibrium

One of the main reasons Fictitious Play was introduced in the 50's is that it hoped that it would be a way for players to converge to a Nash equilibria (or to the set of Nash equilibria).

For zero-sum games this is indeed the case. Indeed, the argument given in the previous example generalises to:

Theorem 5. Assume that (A, B) is a zero-sum game. Then the best response dynamics and also the FP dynamics converges to the set of Nash equilibria of the game.

In fact, one has convergence to Nash equilibria for 2×2 and $2 \times n$ games and several other classes of games.

Example 31. For general 2×2 there are only 4 types of dynamics (up to re-labelling the axis), see the figure below. For example, when $\begin{pmatrix} (1,1) & (0,0) \\ (0,0) & (1,1) \end{pmatrix}$ there are three Nash equilibrium, namely $E = (1/2, 1/2) \times (1/2, 1/2)$ and $(0, 0)$ and $(1, 1)$. The orbits then are as in figure 3 below.

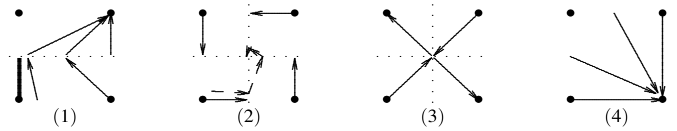


Figure 1: The possible motions in 2×2 games (up to relabeling, and shifting the indifference lines (drawn in dotted lines).

However, in general one does not have convergence.

Example 32. Take

$$A_\beta = \begin{pmatrix} 1 & 0 & \beta \\ \beta & 1 & 0 \\ 0 & \beta & 1 \end{pmatrix} \quad B_\beta = \begin{pmatrix} -\beta & 1 & 0 \\ 0 & -\beta & 1 \\ 1 & 0 & -\beta \end{pmatrix}, \quad (15)$$

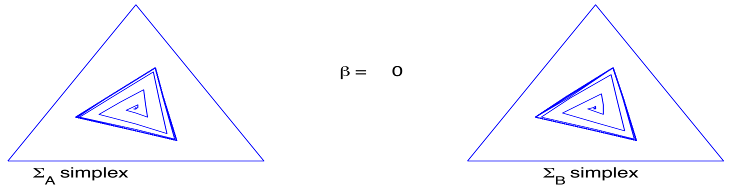
where we use the 2-nd convention.

Note that $E^A \times A^B$ where $E^A := (1/3, 1/3, 1/3)$ and $E^B := (1/3, 1/3, 1/3)'$ is the Nash equilibrium. (How can one work out that there are no other Nash equilibria?).

For $\beta = 0$ this corresponds to the situation that $A = Id$ so wants to copy what player two is doing, and B prefers 3, 2, 1 when player A plays 1, 3 and 2, so player B want so do something different from player A. This games was introduced by the Nobel prize winner Shapley in 1964, to show that the dynamics of FP does not necessarily converge to a Nash equilibrium, but to a periodic orbit.

For $\beta = \phi$ where ϕ is the *golden number* (i.e. $\phi := (\sqrt{5} - 1)/2 \approx 0.618$), the game is equivalent to a zero-sume game (rescaling B to $\tilde{B} = \phi(B - 1)$ gives $A + \tilde{B} = 0$). Hence in this case by Theorem5 play always converges to the interior equilibrium E^A, E^B.

For $\beta \in (\phi, \tau)$ where $\tau \approx 0.915$ the dynamics is chaotic.



5 Fictitious play: a learning model

There are several models for learning, for example

- fictitious play (many people, starting with Brown and Robinson in the 50's), Fudenberg, Levine,....
- reinforcement learning (Roth, Erev, Arthur...)
- no-regret learning (Hart, Mas-Colell, Foster, Young, Kalai, Lehrer,....

Some of these aim to model human behaviour while others are aimed at providing efficient algorithms for computing various generalisations of the Nash equilibrium.

5.1 Best response and fictitious play

Let $x(t)$ and $y(t)$ be the actions (past)play of the two players, and let

$$p(s) = \frac{1}{s}\int_0^s x(u)\,du \text{ and } q(s) = \frac{1}{s}\int_0^t y(u)\,du.$$

Then a player decides to play the following action at time s:

$$x(s) \in \mathcal{B}R_A(q(s)) \text{ and } y(s) \in \mathcal{B}R_B(p(s)) \text{ for } s \geq 1.$$

Note that this means that this equivalent to

$$\begin{aligned} \dot{p}(s) &= \frac{1}{s}(\mathcal{B}R_A(q(s)) - p(s)) \\ \dot{q}(s) &= \frac{1}{s}(\mathcal{B}R_B(p(s)) - q(s)). \end{aligned} \tag{16}$$

Note that if we take the time-reparametrisation $s = e^t$, then we obtain the best-response.

$$\begin{aligned} \dot{p}(t) &= (\mathcal{B}R_A(q(t)) - p(t)) \\ \dot{q}(t) &= (\mathcal{B}R_B(p(t)) - q(t)). \end{aligned}$$

We further denote the maximal-payoff functions

$$\bar{A}(q) := \max_{\bar{p}\in\Delta} \bar{p}Aq \quad \text{and} \quad \bar{B}(p) := \max_{\bar{q}\in\Delta} pB\bar{q}, \tag{17}$$

Let us show that playing fictitious dynamics leads to no-regret.

5.2 The no-regret set

Assume that players A and B have respectively m and n actions.

Definition. A joint probability distribution $P = (p_{ij})$ over $S := \{1, \ldots, m\} \times \{1, \ldots, n\}$ is a *coarse correlated equilibrium (CCE)* for the bimatrix game (A, B) if (p_{ij}), $i = 1, \ldots, m$ and $j = 1, \ldots, n$ is a matrix with all entries ≥ 0 and so that $\sum_{ij} p_{ij} = 1$

$$\sum_{i,j} a_{i'j}p_{ij} \leq \sum_{i,j} a_{ij}p_{ij}$$

and

$$\sum_{i,j} b_{ij'}p_{ij} \leq \sum_{i,j} b_{ij}p_{ij}$$

for all i', j'. The set of CCE is also called the *Hannan set*.

One way of viewing the concept of CCE is in terms of the notion of *regret*. Let us assume that two players are (repeatedly or continuously) playing a bimatrix game (A, B), and let $P(t) = (p_{ij}(t))$ be the empirical joint distribution of their past play through time t, that is, $p_{ij}(t)$ represents the fraction of time of the strategy profile (i, j) along their play through time t. Then $\sum_{i,j} a_{ij}p_{ij}(t)$ and $\sum_{i,j} b_{ij}p_{ij}(t)$ are the players' average payoffs in their play through time t.

For $x \in \mathbb{R}$, let $[x]_+$ denote the positive part of x: $[x]_+ = x$ if $x > 0$, and $[x]_+ = 0$ otherwise. Then the expression

$$\left[\sum_{i,j} a_{i'j}p_{ij}(t) - \sum_{i,j} a_{ij}p_{ij}(t)\right]_+$$

can be interpreted as the regret of the first player from not having played action i' throughout the entire past history of play. It is (the positive part of) the difference between player A's actual past payoff and the payoff she would have received if she always played i', given that player B would have played the same way as she did. Similarly,

$$[\sum_{i,j} b_{ij'} p_{ij}(t) - \sum_{i,j} b_{ij} p_{ij}(t)]_+$$

is the regret of the second player from not having played j'. This regret notion is sometimes called *unconditional* or *external regret* to distinguish it from the *internal* or *conditional regret*[1]. In this context the set of CCE can be interpreted as the set of joint probability distributions with no regret (i.e. the regret is ≤ 0).

5.3 Fictitious play converges to the no-regret set

We now show that continuous-time FP converges to a subset of CCE, namely the subset for which equality holds for at least one i', j' in (18).

Theorem 6. Every trajectory of FP dynamics (16) in a bimatrix game (A, B) converges to a subset of the set of CCE, the set of joint probability distributions $P = (p_{ij})$ over $S^A \times S^B$ such that for all $(i', j') \in S^A \times S^B$

$$\sum_{i,j} a_{i'j} p_{ij} \leq \sum_{i,j} a_{ij} p_{ij} \quad \text{and} \quad \sum_{i,j} b_{ij'} p_{ij} \leq \sum_{i,j} b_{ij} p_{ij}, \tag{18}$$

where *equality* holds *for at least one* $(i', j') \in S^A \times S^B$. In other words, FP dynamics asymptotically leads to no regret for both players.

[1] Conditional regret is the regret from not having played an action i' whenever a certain action i has been played, that is, $[\sum_j a_{i'j} p_{ij} - \sum_j a_{ij} p_{ij}]_+$ for some fixed $i \in S^A$.

Note that an FP orbit $(p(t), q(t))$, $t \geq 1$, gives rise to a joint probability distribution $P(t) = (p_{ij}(t))$ via

$$p_{ij}(t) = \frac{1}{t} \int_0^t x_i(s) y_j(s) ds.$$

When we say that FP converges to a certain set of joint probability distributions, we mean that $P(t)$ obtained this way converges to this set.

Proof of Theorem 6. Let $\bar{A}$ and $\bar{B}$ be defined as in (17).

By the envelope theorem we have that

$$\frac{d\bar{A}(q(t))}{dt} = \bar{p} \cdot A \cdot \frac{dq}{dt}.$$

when $\mathcal{B}R_A(q)$ is unique and $\bar{p} \in \mathcal{B}R_A(q)$. Therefore, since $x(t) \in \mathcal{B}R_A(q(t))$ and $y(t) \in \mathcal{B}R_B(p(t))$ for $t \geq 1$,

$$\frac{d}{dt}\left(t\bar{A}(q(t))\right) = \bar{A}(q(t)) + t\frac{d}{dt}\left(\bar{A}(q(t))\right) = \bar{A}(q(t)) + t \cdot x(t) \cdot A \cdot \frac{dq(t)}{dt}.$$

Using (16) and $\bar{A}(q(t)) = x(t) \cdot A \cdot q(t)$, it follows that

$$\frac{d}{dt}\left(t\bar{A}(q(t))\right) = \bar{A}(q(t)) + x(t) \cdot A \cdot (y(t) - q(t)) = x(t) \cdot A \cdot y(t)$$

for $t \geq 1$. We conclude that for $T > 1$,

$$\int_1^T x(t) \cdot A \cdot y(t)\, dt = T\bar{A}(q(T)) - \bar{A}(q(1)),$$

and therefore

$$\lim_{T \to \infty} \left(\frac{1}{T} \left(\int_0^T x(t) \cdot A \cdot y(t)\, dt \right) - \bar{A}(q(T)) \right) = 0.$$

Note that

$$\frac{1}{T} \int_0^T x(t) \cdot A \cdot y(t)\, dt = \sum_{i,j} a_{ij} p_{ij}(T),$$

(The envelope theorem asserts that, under some examples, the maximum of a function depending on a parameter depends smoothly on this parameter. In this case the parameter is $q(t)$. Can you prove this directly?

Example: $q(t) = \begin{pmatrix} t \\ 1-t \end{pmatrix}$, $A = I$. Then $\bar{A}(q(t)) = \max(t, 1-t)$ and $\bar{p}A \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is equal to 1 and -1

where $P(T) = (p_{ij}(T))$ is the empirical joint distribution of the two players' play through time T. On the other hand,

$$\bar{A}(q(T)) = \max_{i'} \sum_j a_{i'j} q_j(T) = \max_{i'} \sum_{i,j} a_{i'j} p_{ij}(T).$$

Hence,

$$\lim_{T\to\infty} \left(\sum_{i,j} a_{ij} p_{ij}(T) - \max_{i'} \sum_{i,j} a_{i'j} p_{ij}(T) \right) = 0.$$

By a similar calculation for B, we obtain

$$\lim_{T\to\infty} \left(\sum_{i,j} b_{ij} p_{ij}(T) - \max_{j'} \sum_{i,j} b_{ij'} p_{ij}(T) \right) = 0.$$

It follows that any FP orbit converges to the set of CCE. Moreover, these equalities imply that for a sequence $t_k \to \infty$ so that $p_{ij}(t_k)$ converges, there exist i', j' so that $\sum_{i,j}(a_{ij} - a_{i'j}) p_{ij}(t_k) \to 0$ and $\sum_{i,j}(b_{ij} - b_{ij'}) p_{ij}(t_k) \to 0$ as $k \to \infty$, proving convergence to the claimed subset. $\square$

Let us denote the average payoffs through time T along an FP orbit as

$$\hat{u}^A(T) = \frac{1}{T} \int_0^T x(t) \cdot A \cdot y(t)\, dt \quad \text{and} \quad \hat{u}^B(T) = \frac{1}{T} \int_0^T x(t) \cdot B \cdot y(t)\, dt.$$

As a corollary to the proof of the previous theorem we get the following

Proposition 2. In any bimatrix game, along every orbit of FP dynamics we have

$$\lim_{T\to\infty} \left(\hat{u}^A(T) - \bar{A}(q(T)) \right) = \lim_{T\to\infty} \left(\hat{u}^B(T) - \bar{B}(p(T)) \right) = 0.$$

where as before

$$\bar{A}(q) := \max_{\bar{p}\in\Delta} \bar{p} A q \quad \text{and} \quad \bar{B}(p) := \max_{\bar{q}\in\Delta} p B \bar{q},$$

Another consequence is:

Proposition 3. Let (A, B) be a bimatrix game with unique, completely mixed Nash equilibrium (E_A, E_B). If $\bar{A}(q) \geq \bar{A}(E_B)$ and $\bar{B}(p) \geq \bar{B}(E_A)$ for all $(p, q) \in \Sigma$, then asymptotically the average payoff along FP orbits is greater than or equal to the Nash equilibrium payoff (for both players).

5.4 FP orbits often give better payoff than Nash

Consider the one-parameter family of 3×3 bimatrix games (A_β, B_β), $\beta \in (0, 1)$, given by

$$A_\beta = \begin{pmatrix} 1 & 0 & \beta \\ \beta & 1 & 0 \\ 0 & \beta & 1 \end{pmatrix}, \qquad B_\beta = \begin{pmatrix} -\beta & 1 & 0 \\ 0 & -\beta & 1 \\ 1 & 0 & -\beta \end{pmatrix}. \tag{19}$$

This family can be viewed as a generalisation of Shapley's game. This system has been shown to give rise to a very rich chaotic dynamics with many unusual and remarkable dynamical features. The game has a unique, completely mixed Nash equilibrium E, where $E = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \times (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, which yields the respective payoffs

$$u^A(E) = \frac{1+\beta}{3} \quad \text{and} \quad u^B(E) = \frac{1-\beta}{3}.$$

To check the hypothesis of Proposition 3, let $q = (q_1, q_2, q_3)^\top \in \Delta_B$, then

$$\begin{aligned} \bar{A}(q) &= \max\{q_1 + \beta q_3, q_2 + \beta q_1, q_3 + \beta q_2\} \\ &\geq \frac{1}{3}((q_1 + \beta q_3) + (q_2 + \beta q_1) + (q_3 + \beta q_2)) \\ &= \frac{1}{3}(q_1 + q_2 + q_3)(1 + \beta) \\ &= \frac{1+\beta}{3} \\ &= u^A(E) = \bar{A}(E_B). \end{aligned}$$

Moreover, equality holds if and only if

$$q_1 + \beta q_3 = q_2 + \beta q_1 = q_3 + \beta q_2,$$

which is equivalent to $q_1 = q_2 = q_3$, that is, $q = E_B$. We conclude that $\bar{A}(q) > \bar{A}(E_B)$ for all $q \in \Delta_B \setminus \{E_B\}$, and by a similar calculation, $\bar{B}(p) > \bar{B}(E_A)$ for all $p \in \Delta_A \setminus \{E_A\}$. As a corollary to Proposition 3 we get the following result.

Theorem 7. Consider the one-parameter family of bimatrix games (A_β, B_β) in (19) for $\beta \in (0,1)$. Then any (non-stationary) FP orbit Pareto dominates constant Nash equilibrium play in the long run, that is, for large times t we have

$$\hat{u}^A(t) > u^A(E) \quad \text{and} \quad \hat{u}^B(t) > u^B(E).$$

In fact, for this game FP also improves on the set of correlated equilibria. Here we say, that a joint probability distribution $P = (p_{ij})$ is a *correlated equilibrium (CE)* for the bimatrix game (A, B) if

$$\sum_k a_{i'k} p_{ik} \leq \sum_k a_{ik} p_{ik} \quad \text{and} \quad \sum_l b_{lj'} p_{lj} \leq \sum_k b_{lj} p_{lj}$$

for all i, i' and j, j'.

5.5 A conjecture

There are certainly examples of games where the opposite holds, namely where a FP orbit is Pareto dominated by the Nash payoff. However, a numerical study suggests this is extremely rare. For many games FP orbits Pareto dominate Nash play, and conjecturally, for a very large proportion (say %99 percent), FP orbits dominate Nash play for large periods of time.

Note that this all depends on the choice of the matrices.

Theorem 8. Let (A, B) be an $n \times n$ bimatrix game with unique, completely mixed Nash equilibrium E. Then there exists a linearly equivalent game (A', B'), for which $\bar{A}'(q) > \bar{A}'(E_B)$ and $\bar{B}'(p) > \bar{B}'(E_A)$ for all $p \neq E_A$ and $q \neq E_B$, and so for (A', B') FP payoff Pareto dominates Nash payoff.

5.6 Discrete fictitious dynamics

Sometimes it is more natural to consider discrete time, so assume that $t \in \mathbb{N}$. In this case we let $p(0), q(0)$ be the a priori believe at time $t = 0$ of the probability that player B resp A thinks the strategies will be played. The updating rule about these believes is then

$$p(n+1) = \frac{np(n) + e_i}{n+1}, q(n+1) = \frac{nq(n) + e_j}{n+1}$$

So

$$p(n+1) - p(n) = \frac{1}{n}(e_i - p(n)), q(n+1) = q(n) + \frac{1}{n}(e_j - q(n)).$$

This should be considered as the discrete approximation of the continuous best response dynamics

$$\dot{p} = \mathcal{B}R_A(q) - p, \dot{q} = \mathcal{B}R_B(p) - q.$$

6 Reinforcement learning

In this section I will describe the the Arthur and Erev-Roth models for **reinforcement learning**. These are closely related to fictitious play.

Assume:

- at each time period t, each of the two players chooses an action $x(t)$ resp. $y(t)$ (here we will write x^t and y^t for simplicity). In fact, y^t could also be 'nature' or a player which has a totally different way of choosing strategies.
- the payoff for player A is given by a function $u^t = u(x^t, y^t)$ which for pure actions can be written in the form $x^t \cdot Ay^t$. In this current set-up we assume that the payoff is always strictly positive.

Define a variable $\theta_x^t \geq 0$ which describes the 'propensity' of player A to play x at time t which is updated in some manner according to how "good playing x" has been. Let θ^t be the vector of these numbers with x-th component equal to θ_x^t.

Then, at time t, A plays x with probability

$$p_x^t = \frac{\theta_x^t}{\sum_{x' \in X} \theta_{x'}^t}$$

and let p^t be the corresponding probability vector. So the action x_t is chosen according to the probability vector p^t. Several updating rules have been proposed for θ_x^t:

Let $u^t = u(x^t, y^t) \in \mathbb{R}$ be the payoff of player A at time t and e^t be the vector with the x^t component 1 and all other components 0.

(I) Erev-Roth **Cumulative payoff matching (CPM)** (dating back to 1995) is:

$$\theta^{t+1} = \theta^t + u^t e^t.$$

(II) The Athur model from (1993), is closely related:

$$\theta^{t+1} = (\theta^t + u^t e^t)\frac{C(t+1)}{Ct + u^t},$$

where C is the sum of the coordinates of θ^1.

Note that the models are quite similar, the latter is just a rescaled version of the former one. Both models have in common that they reinforce playing a particular action depending on the payoff it resulted in.

In the 2nd model, we have by induction that the sum of the coordinates of θ^t is equal to tC for all $t \geq 1$. Indeed, by the induction assumption, the sum of the coordinates of $\theta^t + u^t e^t$ is equal to $Ct + u^t$, and so the induction step follows. This property makes the 2nd one slighlty easier to work with and in the rest of this section, we will consider only the 2nd model.

Note that player A does not need to observe the actions of player B to determine θ^t, only his own utility pay-off. It is assumed that each pay-off is > 0 and that θ^1 has all strictly positive coordinates. This implies that the probability of choosing action x at time $t+1$ is at least

$$\frac{\theta^1_x}{|\theta^1| + tK}$$

where θ^1 is the sum of the coordinates of the initial propensity vector and K an upper bound for the utility of all actions. It follows that all actions are chosen with positive probability infinitely many times.

Note that in the 2nd model

$$\begin{aligned} p^{t+1} &:= \frac{\theta^{t+1}}{(t+1)C} \\ &= p^t + \frac{u^t}{Ct + u^t}(e^t - p^t) \\ &= p^t + \frac{u^t}{Ct}(e^t - p^t) + \epsilon^t \end{aligned}$$

where $\epsilon^t = O(1/t^2)$. Note that u^t, e^t, ϵ^t are random variables which depends on the actions $x^1, \dots, x^t$ and $y^1, \dots, y^t$ chosen. If we set

$$f(p^t) = E(u^t(e^t - p^t)|\{(x^1, y^1), \dots, (x^t, y^t)\})$$

then we can write the previous equation as

$$p^{t+1} = p^t + \frac{1}{Ct} f(p^t) + \frac{1}{Ct}\mu(p^t) + \epsilon^t$$

where $E(\mu(p^t)|\{(x^1, y^1), \dots, (x^t, y^t)\}) = 0$. Note that p^{t+1} and $f(p^t)$ depends on $x^1, \dots, x^t$ and $y^1, \dots, y^t$.

6.1 A two player version of this each with two actions:

Now do the same for the other player. Then the action y^t player B chooses depends on q^t, and so we obtain the discrete time stochastic process:

$$\begin{aligned} p^{t+1} &= p^t + \frac{1}{Ct} f(p^t, q^t) + \frac{1}{Ct}\mu(p^t) + \epsilon^t \\ q^{t+1} &= q^t + \frac{1}{Ct} g(q^t, q^t) + \frac{1}{Ct}\zeta(q^t) + \epsilon^t. \end{aligned} \tag{20}$$

Note that this is the Euler approximation of a differential equation with decreasing time steps. Indeed, then the points

$$(p^1, q^1), (p^2, q^2), \dots, (p^n, q^n)$$

should correspond to an approximation of the solution of the differential equation

$$\dot{p} = f(p^t, q^t), \dot{q} = g(p^t, q^t)$$

at time

$$\frac{1}{C} + \frac{1}{2C} + \dots + \frac{1}{(n-1)C}.$$

Let us simplicity assume that each of the players has two actions and that the payoff matrices are A and B are equal to $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ where we use the 1st convention. Note that

$$\begin{aligned} f(p^t, q^t) &= E(\frac{u^t(e^t - p^t}{t}) \\ &= p^t(a_{11}q^t + a_{12}(1-q^t))(1-p^t) + \\ &\quad\quad (1-p^t)(a_{21}q^t + a_{22}(1-q^t)(0-p^t) \\ &= p^t(1-p^t)\left((a_{12}-a_{22}) - q^t((a_{21}-a_{11}) + (a_{12}-a_{22}))\right) \end{aligned}$$

and similarly for g. So we get

$$\begin{aligned} f(p,q) &= p(1-p)[\alpha_1 - q(\alpha_1+\alpha_2)] \\ g(p,q) &= q(1-q)[\beta_1 - p(\beta_1+\beta_2)] \end{aligned}$$

where

$$\begin{aligned} \alpha_1 = a_{12} - a_{22}, &\quad \alpha_2 = a_{21} - a_{11} \\ \beta_1 = b_{12} - b_{22}, &\quad \beta_2 = b_{21} - b_{11} \end{aligned}$$

Now compare this with the replicator dynamics

$$\begin{aligned} \dot{p}_i &= p_i[(Aq)_i \; - p \cdot Aq] \\ \dot{q}_j &= q_j[(Bp)_j \; - q \cdot Bp] \end{aligned}$$

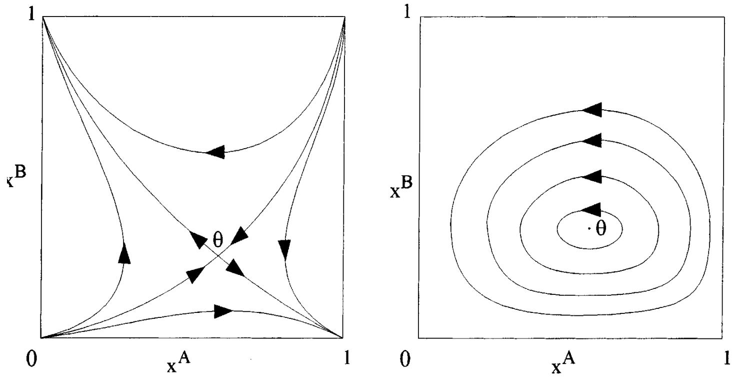
This also gives

$$\begin{aligned} \dot{p}_1 &= p_1[a_{11}q_1 + a_{12}q_2 - p_1(a_{11}q_1 + a_{12}q_2) \\ &\quad -p_2(a_{21}q_1 + a_{22}q_2)] \\ &= p_1(1-p_1)[\alpha_1 - q_1(\alpha_1+\alpha_2)] \\ \dot{q}_1 &= q_1(1-q_1)[\beta_1 - p_1(\beta_1+\beta_2)] \end{aligned}$$

So

$$\dot{p}_1 = f(p_1, q_1), \dot{q}_1 = g(p_1, q_1)$$

is the two person replicator dynamics that we already encountered in Subsection 4.1. As we saw there, the dynamics of this two person replicator system can be completely described. If there is an interior NE then there are two possibilities, where the diagram on the right corresponds to a a game which is equivalent to a zero-sum game.



6.2 Reinforcement learning and replicator dynamics

Based on this, Posch (1997) showed the following:

Theorem 9. In a two-player two strategy game, we have the following possibilities:

- if the game has no strict Nash equilibrium and equivalent to a zero sum game (as in the previous figure on the right), then the learning algorithm has a continuum of asymptotically cycling paths. Almost all paths that are not asymptotically cycling converge either to the interior fixed point or to the boundary;

- at least one strict Nash equilibrium and $C \geq a_{jk}, b_{jk}$ for $j, k = 1, 2$, then the learning algorithm a.s. converges to the set of strict Nash equilibria. All strict Nash equilibria are attained in the limit with positive probability.

6.3 Stochastic approximation

To prove this, note that the abstract form of the stochastic process is

$$x_{n+1} - x_n = \gamma_{n+1} V_{n+1}$$

where V_{n+1} is a random variable of the form

$$V_{n+1} = F(x) + U_{n+1}$$

and where U_{n+1} is a random variable. Since $\gamma_n \to 0$, this should be related to the differential equation

$$\frac{dx}{dt} = F(x).$$

There is an extensive literature about this, see the introductory section.

In these notes we will not be able to discuss these results, and more specifically to what extend the learning process from equation (20) can indeed be modelled by the replicator differential equation. To explain the issues that are at stake, let us show specifically what's going on.

6.4 What happens if C is not large enough?

Proposition 4. Suppose that $0 < C < a_{k,l}, b_{k,l}$ for all k, l. Then

$$Prob\{\lim_{t\to\infty} p^t \to 1, \lim_{t\to\infty} q^t \to 1\} > 0.$$

Proof. Let us show that there is a positive probability that $p^t \to 1$. To do this, let us show below that *if player A plays action 1 forever* then $\prod_{t=1}^{\infty} p^t > 0$. This in turn implies that player A chooses action 1 forever with positive probability, because after all $\prod p^t$ is the probability that player A chooses the 1st action forever. Moreover, as $p_t < 1$ and $\prod_{t=1}^{\infty} p^t > 0$ implies $p^t \to 1$. (Another way of seeing this is that the propensity θ_x^t for the first action $x = 1$ keeps increasing, while θ_2^t remains constant. This then implies that with positive probability $p^t \to 1$.)

Note that $\prod_{t=1}^{\infty} p^t > 0$ is equivalent to $\sum(1 - p^t) < \infty$. Note that

$$p^{t+1} = p^t + \frac{u^t}{Ct + u^t}(e^t - p^t).$$

This means that *if player A chooses action* 1 at time t and player B action j then

$$p^{t+1} = p^t + \frac{a_{1j}}{Ct + a_{1j}}(1 - p^t).$$

So writing $d^t = 1 - p^t$ we get

$$d^{t+1} = d^t - \frac{a_{1j}}{Ct + a_{1j}} d_t$$

or

$$\frac{d^{t+1}}{d^t} = 1 - \frac{a_{1j}}{Ct + a_{1j}} = 1 - \frac{a_{1j}}{Ct} + O(1/t^2).$$

Since $a_{ij} > C$ for all i, j, there exists $\alpha > 1$ and t_0 so that for $t \geq t_0$

$$\frac{d_{t+1}}{d_t} < 1 - \frac{\alpha}{t}.$$

This implies by the Raabe test that $\sum_{t=1}^{\infty} d_t$ converges. (The Raabe test states the following. Assume $|c_n/c_{n+1}| \to 1$ and $n(|c_n/c_{n+1}| - 1) \to R$. Then $\sum c_n$ converges if $R > 1$ and diverges if $R < 1$.) □

6.5 The dynamics of this learning process

Theorem 10. In the case of a 2×2 zero sum case, the learning algorithm has a continuum of asymptotic cycling paths.

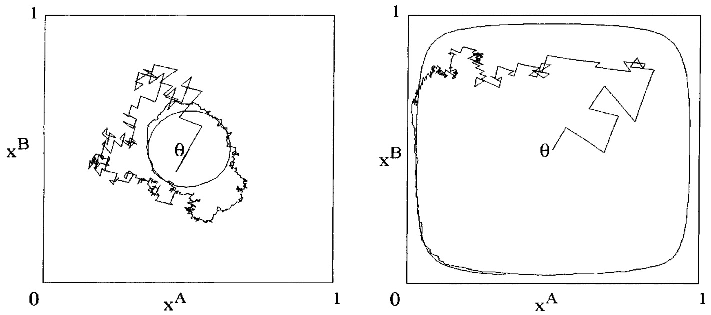


Fig. 2. Two runs of the cycling learning process for a constant sum game. The interior fixed point (the Nash equilibrium) is $\boldsymbol{\theta} = (\frac{1}{2}, \frac{1}{2})$, from where we started the run

Similarly, we have

Theorem 11. Assume that the game has at least one strict Nash equilibrium. If $C \geq a_{jk}, b_{jk}$ then the learning algortithm a.s. converges to the set of strict Nash equilibria. Moreover, each Nash equilibrium is attained with positive probability.

Proof. There are several versions of this result. For example Posch (1997), Hopkins & Posch (2005) and references in these papers. □

6.6 What if the opponent has a different strategy: computer experiments

Consider the following game: $\begin{pmatrix} 2,1 & 1,2 \\ 1,2 & 2,1 \end{pmatrix}$.

Suppose the 2nd player uses

1. fictitious play;
2. takes a (myopic) best response to player 1's current action;
3. plays the minmax strategy.

Then player 1's average payoff converges rapidly to 1.5. Indeed, Beggs [2005] did some computer simulations. Against each opponent the ER rule was run 100 times in a run of length 10,000, with initial reinforcements (1,1.5).

The mean average payoff was

1. 1.48 st.dev. 0.04

2. 1.49 st.dev. 0.01

3. 1.5 st.dev. 0.003

7 No regret learning

There are several results that describe 'no regret learning' algorithms to 'learn' the correlated equilibrium (CE) of a bimatrix game.

To put this in context, we have the set of Nash equilibria (NE), the correlated equilibria (CE) and the course correlated equilibria (CCE). These are related as follows:

$$NE \subset CE \subset CCE.$$

We saw that the best response dynamics converges to the CCE set.

Let us first recall the definition of the the CE set. Assume that A has m actions and player B has n actions. We say that the matrix (p_{ij}), $i = 1, \ldots, m$ and $j = 1, \ldots, n$ is a probability distribution if all its entries are ≥ 0 and $\sum_{ij} p_{ij} = 1$. A joint distribution is a *correlated equilibrium (CE)* for the bimatrix game (A, B) if

$$\sum_k a_{i'k} p_{ik} \leq \sum_k a_{ik} p_{ik} \quad \text{and} \quad \sum_l b_{lj'} p_{lj} \leq \sum_k b_{lj} p_{lj}$$

for all i, i' and j, j'.

So this means that if you consider p_{ij} as the proportion of time up to time t that action i, j was chosen, then $t\left(\sum_k a_{ik} p_{ik}\right)$ is the payoff resulting from action i. The first inequality means that player A would not have been better off by switching action i to action i'. The 2nd inequality means that the same holds for player j.

Note that a Nash equilibrium corresponds to the special case where (p_{ij}) is a product distribution, so correspond to the situation that there are two probability vectors p^*, q^* and that $p_{ij} = p^*_i \cdot q^*_j$.

In this section we will discuss some results related to learning models which converge to CE.

7.1 Hart and Mas-Colell's regret matching

Suppose that the two players have played actions x^i, y^i for time $i = 1, \ldots, t$. Let $WA^i(j,k)$ and $WB^i(j,k)$ be the regret at time i that the player chose j instead of k. More precisely, for $i = 1, 2, \ldots, t$, define

$$WA^i(j,k) = \begin{cases} e_k \cdot Ay^i & \text{if } x^i = e_j \\ x^i \cdot Ay^i & \text{if } x^i \neq e_j \end{cases}$$

and similarly for WB. So this gives the payoff A would have received at time i, assuming player B would have done the same, if only he had played k whenever he actually played j. Then define

$$DA^t(j,k) = \frac{1}{t}\left(\sum_{i=1}^{t}[WA^i(j,k) - x^i \cdot Ay^i]\right).$$

So this is what player A would have gained (or lost) on average up to time t had he played action k whenever he actually played j. Now define

$$RA^t(j,k) = \max(DA^t(j,k), 0).$$

Let j^* be the action of player A at time t and define the component of p^{t+1} by

$$\begin{array}{lll} p_j^{t+1} & = \dfrac{1}{\mu}RA^t(j^*, j) & \text{for all } j \neq j^* \\ p_{j^*}^{t+1} & = 1 - \sum_{j \neq j^*} p_j^{t+1} & \text{when } j = j^* \end{array}$$

Here μ is chosen so large that the above vector is a probability vector. This means that the probability of switching to a different stratefy is proportional to their regrets relative to the current strategy. For player B define similarly RB^t and q^{t+1}.

Theorem 12 (Hart and Mas-Colell)**.** Provided we fix μ sufficiently large, if player A follows this algorithm then almost surely $RA^t(j,k) \to 0$ as $t \to \infty$.

Moreover,

Theorem 13 (Hart and Mas-Colell)**.** Provided we fix μ sufficiently large, if both players follow this algorithm then the resulting frequency of actions up to time t tends to the CE set as $t \to \infty$.

Foster-Fohra and Fudenberg-Levine have related results.

We will not explain the proofs of these theorems, but prove a related results namely "universal consistency".

7.2 Min-max solutions and zero-sum games

Before going into no regret learning it is good to state a well-known fact which is related to zero-sum games.

Theorem 14. For any matrix A one has

$$v_A := \max_x \min_y x \cdot Ay = \min_y \max_x x \cdot Ay := v_B. \qquad (21)$$

In fact, (21) is equivalent to the existence of a Nash equilibrium (x^*, y^*) of the game $(A, -A)$ and $v_A = v_B = x^* \cdot A \cdot y^*$ is called the value of the zero-sum game.

Proof. Since $\min_y x \cdot Ay \le \min_y \max_x x \cdot Ay = v_B$, we have $v_A \le v_B$. To prove the opposite inequality, let (x^*, y^*) be a Nash equilibrium of the zero-sum game $(A, -A)$. This means $x^* \in BR_A(y^*)$ and $y^* \in BR_B(x^*)$ where $B = -A$. This equivalent to the requirement that for all x, y,

$$x \cdot Ay^* \le x^* \cdot Ay^* \text{ and } x^* \cdot Ay^* \le x^* \cdot Ay. \qquad (22)$$

(Note that $B = -A$ and hence the 2nd inequality is $\le$). The previous two inequalities are equivalent to

$$\max_x x \cdot Ay^* = x^* \cdot Ay^* = \min_y x^* \cdot Ay.$$

It follows that

$$v_B := \min_y \max_x x \cdot Ay \leq \max_x x \cdot Ay^* = x^* \cdot Ay^*$$
$$= \min_y x^* \cdot Ay \leq \max_x \min_y x \cdot Ay := v_A.$$

This proves the first assertion of the theorem. In fact, (21) implies that there exists a Nash equilibrium. Indeed, take x^*, y^* so that $\min_y x^* \cdot Ay = v = \max_x x \cdot Ay^*$. Let us show that x^*, y^* is a NE. Indeed for all x, y,

$$x^* \cdot Ay \geq \min_y x^* \cdot Ay = v \text{ and } x \cdot Ay^* \leq \max_x x \cdot Ay^* = v$$

Substituting for x^*, y^* for x, y it follows that $v = x^* \cdot Ay^*$ and

$$x^* \cdot Ay \geq x^* \cdot Ay^* \geq x \cdot Ay^*.$$

and hence the conditions (22) for NE are satisfied. □

7.3 Blackwell approachability theorem

Assume that player A decides to play p^t, $t = 1, \dots$ and his adversary plays q^t, $t = 1, 2, \dots$. Now assume that the player receives a payoff vector (rather than a payoff number) and denote this payoff $A(p^t, q^t) \in \mathbb{R}^k$ and let $a_t = (1/t)\sum_{t=1}^t A(p_t, q_t)$.

We say that $\mathcal{C}$ is *approachable* if for each probabilities $\{p^i, q^i\}_{i=1}^{t-1}$ there exists a choice p^t (which is independent of the choice for q^t so that a_t converges to a convex set $\mathcal{C} \subset \mathbb{R}^k$ (in the Euclidean norm). Blackwell's Approachability Theorem gives a necessary and sufficient condition for $\mathcal{C} \subset \mathbb{R}^k$ to be approachable. In the setting of this theorem it will turn out that p^t only depends on $\mathcal{C}$, a^{t-1} and $A(p^{t-1}, q^{t-1})$.

Note that $A(p, q)$ can be written as

$$A(p, q) = \sum_{i=1}^n \sum_{j=1}^m p_i A_{ij} q_j$$

but where A_{ij} is a vector.

Theorem 15 (Blackwall's Approchability)**.** For any closed convex set $\mathcal{C}$ the following are equivalent.

1. $\mathcal{C}$ is approachable;
2. for each q there exists p so that $A(p,q) \in \mathcal{C}$;
3. every half space containing $\mathcal{C}$ is approachable.

Proof. (2) $\implies$ (3) Consider a half-space $H = \{a \in \mathbb{R}^k; n \cdot a \leq v\}$ which contains $\mathcal{C}$.

$$\forall q \exists p \text{ with } A(p,q) \in \mathcal{C} \implies \forall q \exists p \text{ with } n \cdot A(p,q) \leq v \implies$$

$$\exists q \forall q \text{ with } n \cdot A(p,q) \leq v \implies \mathcal{C} \text{ is approachable}$$

Here the exchange of $\forall q \exists p$ to $\exists q \forall p$ follows from the minmax theorem and in the conclusion one chooses the $p^t = p$ where p is from the last line.

(3) $\implies$ (2) Since *each* half-space $H \supset \mathcal{C}$ is approachable, there exists for each such half-space H and for each q some p with $A(p,q) \in H$. Since this holds for each such half-space we also have $\forall q\ \exists p$ with $A(p,q) \in \mathcal{C}$.

(1) $\implies$ (3) trivially follows from $\mathcal{C} \subset H$.

(3) $\implies$ (1) is the most interesting part of the proof. Let $P(a_t)$ be the closest point in $\mathcal{C}$ to a_t, let $n_t = P(a_t) - a_t$ and let $v_t = P(a_t) \cdot n_t$. Then let H_t be the half-space containing $\mathcal{C}$ through $P(a_t)$ orthogonal to n_t. That is, $H_t = \{a; n_t a \leq v_t\}$. Since H_t is approachable, $\forall q\ \exists p$ so that $n_t A(p,q) \leq v_t$. By the minmax theorem this implies $\exists p\ \forall q$ so that $n_t A(p,q) \leq v_t$. Let p_t be this choice. With some further work one can show that $||a_t - P(a_t)|| \leq A/\sqrt{t}$ and so $a_t \to \mathcal{C}$. □

7.4 Universal consistency

Let us given an application of this. Take a real valued payoff $A(p,q)$ and consider vector valued $\hat{A}(p,q)$ with components

$A(p,q) - A(e_i, q)$, $i = 1, \ldots, n$. So this is the gain or loss if player A would choose strategy p instead of strategy i.

Now consider the convex region $\mathcal{C} = \{a; a_i \leq 0 \,\forall i\}$. For each q there exists p so that each of the components of $\hat{A}(p,q) \leq 0$: choose $p = e_{i^*}$ where $i^* = \arg\min A(i,q)$. It follows that the 2nd condition of Blackwell's approachability theorem is satisfied. In particular there exists a strategy p^t so that for each $q^1, q^2, \ldots, q^t$ one has

$$\limsup_{t\to\infty} \left(\frac{1}{t} \sum_{s=1}^{t} A(p^s, q^s) - \min \frac{1}{t} \sum_{s=1}^{t} A(i, q^s) \right) \leq 0.$$