Multi-Agent Systems: Homework Assignment 1

Konstantin Mihhailov, Haohui Zhang, Roel van de Laar, Henrik Kathmann

October 17, 2022

1 Game Theory: Concepts and Nash Equilibrium

1.1 Iterated Best response dynamics for a simple matrix game

Consider the following simultaneous two player game: Player 1 (row player) can play U, M or D, while player 2 (column player) can pick L,C or R. Their pay-off is given in figure 1.

- Are there any dominated strategies?

A dominated strategy is a strategy that is worse in all circumstances, compared to the other strategies a player can play. If we look at figure 1, we can see that the column player can play L, C or R. The possible utilities the column player gets from playing one of these moves are the right values in each column (the left values correspond to the utilities for the row player). We would say that one of these moves is dominated if all possible values in a respective column, lets say L (3, 2, -3), are worse than all values in column C (0, 6, 0) or R (2, 1, -1). The same goes for the row player, but for the row player we compare the left values of each row to the left values of all other rows. Using this logic, we can determine that for both the column and row player there are no dominated strategies.

- The concept of Nash equilibrium plays a central role in game theory and will be discussed at length in the next assignment(s). But some basic intuition can be gleaned from the following approach (known as: best response dynamics). One way to visualise Nash equilibria is to interpret the simultaneous game as a repeated sequential one: One player is allowed to make a move, then the other player makes a move, after which the first player can make another move and so on. What happens when you do this for this game?

Assuming that both players are rational and always play the move that provides them with the highest utility, we can determine if there is a NE in the simultaneous two player game. A NE is a theory within game theory where each player has no incentive to change their strategy because their current strategy is just strictly better then all other strategies this player can choose. One way to think about this is if you were able to freeze both time and all players, and then unfreeze the players one by one and ask them if they want to change their strategy (knowing that the other players cannot change their strategy), every player that you choose to unfreeze will keep playing their current strategy.

Now, if we return to our current game we can see that for values at index (M, C) there is a pure Nash equilibrium (mutual best response) because it does not matter where we start the game, if both

	L	C	R
U	$4,3$	$2,0$	$8,2$
M	$8,2$	$7,6$	$-1,1$
D	$6,-3$	$0,0$	$1,-1$

Figure 1: Simultaneous two player game matrix.

		Player 2		
		Rock	Paper	Scissor
Player 1	Rock	0, 0	-1, 1	1, -1
	Paper	1, -1	0, 0	-1, 1
	Scissor	-1, 1	1, -1	0, 0

Figure 2: Rock paper scissors pay-off matrix.

Chicken

	swerve	straight
swerve	$0, 0$	$-1, 1$
straight	$1, -1$	$-5, -5$

Figure 3: Chicken game pay-off matrix.

players are rational and try to maximize their utility, the game will always end up at this position. We can illustrate this with an example: if the row player starts he/she will choose any of the rows because there are no dominated strategies. Therefore, lets say the row player chooses U. Then, the column player will try to maximize its utility and choose L. As a response to this, the row player will also try to maximize its utility and choose M, which results in the column player choosing C and then the row player will choose M again, and so forth. For any other starting row or column we choose, we will always end up at row M and column C if both players try to maximize their output.

- Can you think of simple games where the outcome of this dynamics would be qualitatively different?

One game where the outcome of this dynamics is different, is the game of rock, paper, scissors. In this game, both players have to simultaneously choose between rock, paper and scissors. The resulting pay-off is shown in figure 2. As we can see in the pay-off matrix, if both players decide on the same strategy they tie and get 0 utility. However, if both players play a different strategy, then one player wins the round and gains 1 utility while the other players loses 1 utility. Now, if we consider NE theory, it is clear that there is no pure Nash equilibrium in this game because both players do not want to match each others strategy. Additionally, if we would freeze time again and unfreeze the players one by one, the losing players would decide to change strategy every time (which would keep going back and forth forever). The game of rock, paper and scissors does have a mixed equilibrium, where both players will decide to play every strategy exactly 1/3 of the time.

Another game where the outcome of this dynamic is different, is the game of chicken. The pay-off matrix for this game is shown in figure 3. Similarly to the rock, paper, scissors game, there is no incentive to play the same strategy because in both cases, it is always in the best interest of a player to not play the same strategy as its opponent. Therefore, if we would freeze time again and unfreeze the players one by one, they will always choose to play the opposite strategy of their opponent. Again, in this game there is nu pure NE but a mixed NE.

1.2 Elimination of dominated strategies

A game is played with a group of say 50 players (e.g. all students sitting in a large auditorium). Upon request, they all pick an (integer) number between 1 and 100 (inclusive), and write it down on a card without communicating with any of the other players. These cards are collected and the average (say m) of all the numbers is computed. The winner is the person(s) whose chosen number was closest to 2/3 of the mean m.

- What would be a rational strategy?

To determine what the rational strategy is for this game, certain aspects need to be considered. Namely, the player needs to consider/attempt to predict what the other players will do, then the player will need to adjust their strategy accordingly. Assuming every player is rational, these steps would be key ingredients to achieve Nash Equilibrium as each player would be choosing their optimal response to the actions of other players.

If the main player presumes the average integer, which includes the main players own integer, then the optimal strategy for the main player will be to assume the winning integer would be 2/3 of the average. Following on from this, assuming all players are rational, the average can not be more than 100, so it has to be less. Hence, the optimal strategy for all the players would be to choose an integer below 67 (2/3 of 100). Following on from this, The players would know that since the average can not be more than 67, the optimal winning strategy would be to select an integer 2/3 of 67.

Assuming every player is rational, this process continues in a cascading fashion until it is established that the Nash Equilibrium for this question is for all players to choose the integer of 1.

- If you had to play this game, would you play the above rational strategy? Explain.

If we assume that we play against people that did not encounter this game before and are not purely rational, then no, it would not be optimal to play the above rational strategy as it is very likely that 2/3 of the average would not cascade down to the integer of 1. In reality, most people do not act fully rationally. Hence, above strategy would not be the optimal winning strategy. However, it is safe to assume that repetition and expertise would diminish the non-equilibrium moves. So if we were to play this game repeatedly against same players, and the players understood what is happening through the repetition, it would be wise to adjust the responses closer and closer to the rational Nash equilibrium of 1 as the repetitions progress.

1.3 Cournot's Duopoly

Cournot's duopoly is a game that models economic competition with strategic substitutes. Strategies are called strategic substitutes if an increase in your strategy will cause the competitor to decrease the use of his strategy ("if you increase your part, I will reduce mine"). This contrasts with strategic complements when an increase in the strategy of one player causes the other player to follow suit ("if you increase your part, I will do the same.".)

Two companies make an interchangeable product (e.g. bottled water). Both need to determine (simultaneously!) the quantity they will produce (say for next week). Call these quantities q1 and q2, respectively. The unit price p(price of each unit, e.g. one bottle) of the product in the market depends on the total produced quantity q1+q2. Specifically:

$$p(q1, q2) = \alpha - \beta(q1 + q2)(\alpha, \beta > 0)$$

Firm 1 can produce each unit at a unit-cost c1, whereas the unit-cost for firm 2 equals c2.

- What is the best response for each company given the quantity the other company will produce?

After rewriting the formula, the total revenue for the first firm is as following:

$$Tr = \alpha q1 - \beta q1^2 - \beta q1q2$$

Therefore, the marginal revenue for the first firm is:

$$MR1 = \alpha - 2\beta q1 - \beta q2$$

Now, we set the marginal revenue equal to the marginal costs. Therefore, the best response function for the first firm would be:

$$q1 = (\alpha - \beta q2 - c1/2)\beta$$

For the second firm, we can do the same procedure with the following best response function:

$$q2 = (\alpha - \beta q1 - c2)/2\beta$$

- Suppose the companies need not decide on their quantity at the same time, but can react to one another (an unlimited number of times). What will be the outcome? (Provide a diagram.)

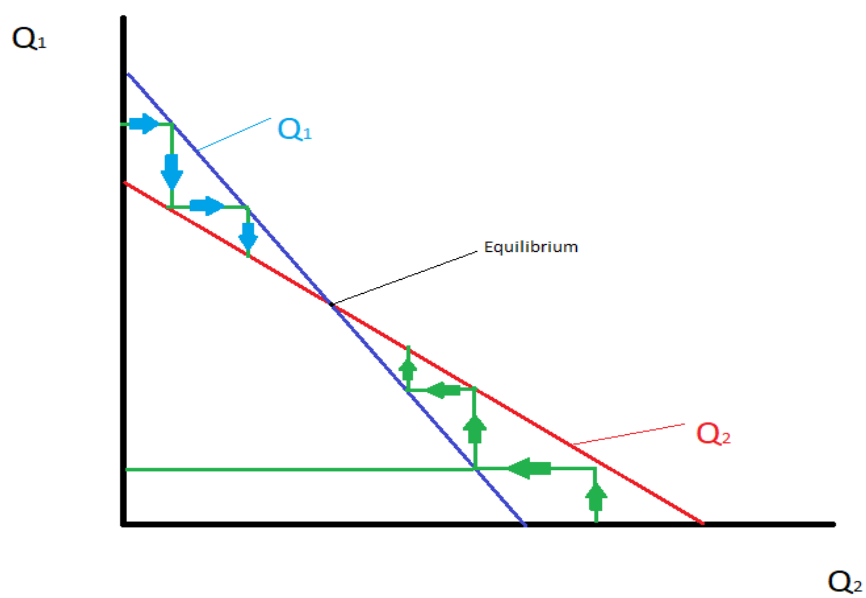
In this case, since the game plays out in a sequential order, one player/firm is leading. Hence, that firm/player chooses a quantity of production. Then, firm/player 2 observes first players quantity and responds with its own quantity of production. Here, the two reaction functions need to be substituted. Substituting BRF1 into BRF2 gives the following function:

$$q1 = (\alpha + c2 - 2c1)/3\beta$$

The other way around one receives the following function:

$$q2 = (\alpha - 2c2 + c1)/3\beta$$

This may be illustrated in the following diagram:



1.4 Ice cream time!

Three competing ice-cream vendors (Alice, Bob and Charlize) are trying to sell their refreshments to tourists on the beach. We are making the following assumptions:

1. the beach has total length of 1, while its width is uniform and much smaller than its length.So the beach can be represented as a line segment of length 1, and each position on the beach can be represented by the position parameter 0 x 1.
2. Tourists are uniformly distributed along the total length of the beach and will buy their ice-cream at the stall that is closest to their location.

1.4.1 Questions.

- On a beautiful summer morning Charlize makes her way to the beach and upon arrival finds that her two competitors have already set up their stalls: Alice at location a = 0.1 and Bob at location b = 0.8. Discuss what Charlize's best response is: i.e. what location should she choose, given a = 0.1 and b = 0.8?

In this situation, since Alice and Bob already set up their stalls, we only need to find Charlize's best response. The best response to strategy profile s_-i is for s_i in S_i,

$$u_i(s_i^*, s_{-i}) >= u_i(s_i, s_-i)$$

Now there is three situations:

1. Charlize chooses the position to the left of Alice. This way she shares no more than 0.1 revenue.
2. Charlize chooses the position to the right of Bob. This way she will share no more than (1-0.8) = 0.2 revenue.
3. Charlize chooses anywhere between Alice and Bob. This way she will get (0.8-0.1)/2 = 0.35 revenue.

0.35 is bigger than 0.2 and 0.1, according to the best response equation, Alice should choose anywhere between Alice and Bob, which is [0.1, 0.8].

- Same question as above, but now assume that all we know is that a = 0.1 and a ¡ b 1.

In this question, we only know a = 0.1. We don't know where Bob will set his stall, it can be anywhere from 0.1 to 1. Thus, to find Charlize's best response, we should suspect Bob's location, and then maximize Charlize's revenue. The first situation is that, if Bob chooses the location which is b $\leq$ 0.1, Charlize should set her stall just next to Alice and the rest 0.9 revenue will be enjoyed by Charlize alone. In all other cases we can eliminate the 0.1 revenue to the left of Alice because the location between 0 and 0.1 is a strictly dominated action. Now, Alice, Bob and Charlize share the rest 0.9 revenue and we need to find the Maximin strategy for Charlize. The Nash equilibrium is that, Alice, Bob and Charlize all have the same 0.3 revenue, which is Bob and Charlize both at the point of 0.7. So the minimum pay-off for Charlize is 0.3. Using this knowledge, we can separate our answer into 4 situations according to Bob's position.

1. If Bob's position is (0 - 0.1), Charlize should choose the position just to the right of Alice, which is infinitely close to 0.1. The pay-off (utility) will be 0.9.
2. If Bob's position is (0.1 - 0.7), Charlize should choose the position just to the right of Bob. The pay-off will be more than 0.3, but less than 0.9.
3. If Bob's position is 0.7, Charlize should choose the position just to the right of Bob or anywhere between 0.1 and 0.7. The resulting pay-off will be 0.3.
4. If Bob's position is (0.7 - 1), Charlize should choose the position just to the left of Bob. The pay-off will be more than 0.3, but less than 0.45.

- Earlier that morning, Bob arrived and discovered that Alice had already set up her stall at a = 0.1, while Charlize hadn't shown up yet. But Bob knows that Charlize will arrive before too long, and that she will try to position herself in such a way as to maximize her revenue. What location should Bob pick in order to maximize his own expected revenue?

This question can also be solved using the Maximin strategy. Similar to the above situation, the location between 0 and 0.1 is a strictly dominated action. Meaning that pay-off can be eliminated. The Nash equilibrium is that Alice, Bob and Charlize divide the remaining 0.9 and each enjoy 0.3 pay-off. The difference is that Charlize will choose a position to maximize her revenue. Thus, it is a maximin situation for Bob. $V_i^{mami} = max_{s_i} min_{s_j} u_i(s_i, s_j) = 0.3$. $S_i^{mami} = argmax_{s_i} min_{s_j} u_i(s_i, s_j) = 0.7$. Bob should put his stall at position 0.7, because the maximin value is 0.3, if Bob puts his stall somewhere else, Charlize can put her stall just next to Bob and make Bob's pay-off less than 0.3.

- At sunrise that morning, Alice arrived before both Bob and Charlize, and set up her stall at location a = 0.1. However she knows for sure that the other two vendors will show up soon. Where should she have set up her stall in order to maximise her expected revenue?

This question can be solved using the Maximin strategy as well. However, the situation is different, both Bob and Charlize will choose a position to maximize their revenue. And the Nash equilibrium here is that Alice, Bob and Charlize divide the 1 and each enjoy 1/3 pay-off. Now, the common answer maybe that Alice chooses either the position at 1/3 or 2/3. However, picking the location at 1/3 in this situation is not the dominated strategy, and we cannot eliminate the part left Alice now. For example, if Alice put her stall at 1/3, and we now find the maximin strategy for Bob. If we eliminate the 1/3 left to Alice, the Nash equilibrium for Bob and Charlize is to divide the left 2/3 into 3 parts, and in this situation, the only choice for Bob is choose the 7/9 point to put his stall, and his maximin value is 2/9, however 2/9 is less than 1/3, so it can be proven that the strategy to pick 1/3 is not a strictly dominated strategy, and Bob will definitely choose the position infinite close to the left of Alice and then Charlize will choose the position infinite close to the right of Alice, and now, Alice's pay-off is infinite close to 0. Thus, the right solution is to find a weakly dominated strategy. Suppose the position Alice choose is x, x must equal to (1-x)/3. After solving, x = $(1 - x)/3$, we get x=0.25, and now Bob will definitely not to choose the position infinite close to the left of 0.25, because if Bob choose the 0.75, he have the opportunity to get a pay-off more than 0.25. So the 0.75 in this situation for Bob is the maximin strategy. And then, no matter where Charlize picks, Alice will get at least 0.25. So the answer is, the best strategy for Alice is choosing the 0.25 or 0.75.