Multi-Agent Systems

Homework Assignment 1

Martynas Vaznonis (2701013)
m.vaznonis@student.vu.nl

1 Game Theory: Optimality Concepts and Nash Equilibrium

1.1 Odd or even game

Questions:

1. Left is the pay-off matrix with p and q probabilities of choosing action 1 for agents A and B respectively. Right is the regret table placed here retrospectively for future convenience.

A \ B	**1** (q)	**2** (1-q)
1 (p)	-2, 2	3, -3
2 (1-p)	3, -3	-4, 4

A \ B	**1** (q)	**2** (1-q)
1 (p)	5, 0	0, 5
2 (1-p)	0, 7	7, 0

2. The regret values for each agent are calculated below according to the formula $R_1(s_1, s_2) = u_1(BR_1(s_2)) - u_1(s_1, s_2)$:

$$R_A(1,1) = 5 \qquad R_A(1,2) = 0 \qquad R_A(2,1) = 0 \qquad R_A(2,2) = 7$$
$$R_B(1,1) = 0 \qquad R_B(1,2) = 7 \qquad R_B(2,1) = 5 \qquad R_B(2,2) = 0$$

For agent A, the maximum regret values for $s_1 = 1$ and $s_1 = 2$ are 5 and 7 respectively. Thus, the best pure strategy for regret minimization is to play 1.

For agent B, the maximum regret values for $s_1 = 1$ and $s_1 = 2$ are 7 and 5 respectively. Thus, the best pure strategy for regret minimization is to play 2.

3. For agent A, the expected regret values for $s_2 = 1$ and $s_2 = 2$ are $5p$ and $7(1-p)$ respectively. Thus, $5p = 7 - 7p \Rightarrow p = \frac{7}{12}$.

For agent B, the expected regret values for $s_2 = 1$ and $s_2 = 2$ are $5(1-q)$ and $7q$ respectively. Thus, $7q = 5 - 5q \Rightarrow q = \frac{5}{12}$.

4. For agent A, the minimum values for $s_1 = 1$ and $s_1 = 2$ are -2 and -4 respectively. Thus, the best pure safety strategy is to play 1.

For agent B, the minimum values for $s_1 = 1$ and $s_1 = 2$ are -3 and -3 respectively. Thus, the best pure safety strategy is to play 1 or 2.

5. For agent A, the expected utility values for $s_2 = 1$ and $s_2 = 2$ are $3(1-p) - 2p = 3 - 5p$ and $3p - 4(1-p) = 7p - 4$ respectively. Thus, $3 - 5p = 7p - 4 \Rightarrow p = \frac{7}{12}$ and expected utility $\mathbb{E}u_A(p, s_2 = 1) = \mathbb{E}u_A(p, s_2 = 2) = 7p - 4 = \frac{1}{12}$ which is higher than the expected utility for pure strategies (-2), making it preferred.

   For agent B, the expected utility values for $s_2 = 1$ and $s_2 = 2$ are $2q - 3(1-q) = 5q - 3$ and $-3q + 4(1-q) = 4 - 7q$ respectively. Thus, $5q - 3 = 4 - 7q \Rightarrow q = \frac{7}{12}$ and expected utility $\mathbb{E}u_B(q, s_2 = 1) = \mathbb{E}u_B(q, s_2 = 2) = 5q - 3 = -\frac{1}{12}$ which is higher than the expected utility for pure strategies (-3), making it preferred.

1.2 Travelers' dilemma: Discrete version

Questions:

1. The pay-off matrix can be seen below with respective probabilities in parentheses.

A \ B	**1** (v)	**2** (w)	**3** (1-v-w)
1 (p)	1, 1	1+a, 1-a	1+a, 1-a
2 (q)	1-a, 1+a	2, 2	2+a, 2-a
3 (1-p-q)	1-a, 1+a	2-a, 2+a	3, 3

2. Because $a > 0$ and $a < 0.5$ the best response with a pure strategy is to always play the same move as the opponent. If the player B's strategy is to play $x \in \{1, 2, 3\}$, player A can either get x, if he also plays x, $x - a$, if he plays $y > x$, or $y + a$ if $y < x$. Because of the bounds on a, it holds that $y + a < x - a < x$. Thus, there are three pure NE strategies, for both agents to play the same number 1, 2, or 3.

3. First, the expected utilities for each of the opponent's actions are set to be equal:

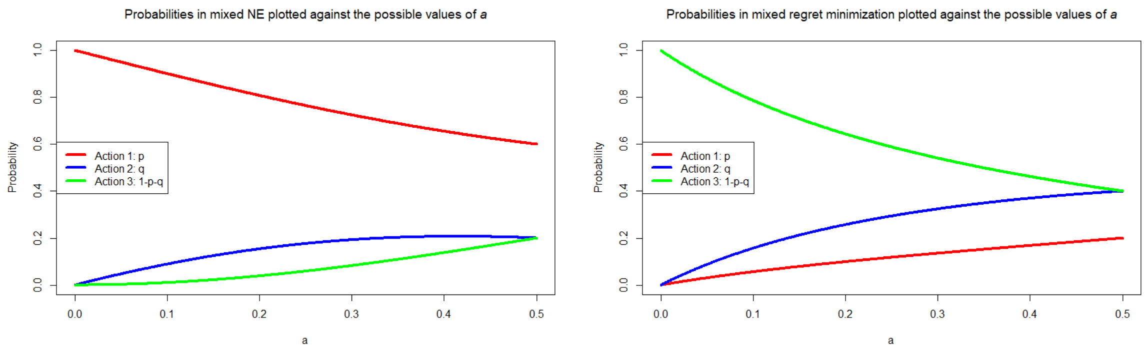
$$p + q(1+a) + (1-p-q)(1+a) = p(1-a) + 2q + (1-p-q)(2+a) \Rightarrow p = \frac{1-aq}{1+a}$$
$$p + q(1+a) + (1-p-q)(1+a) = p(1-a) + q(2-a) + 3(1-p-q) \Rightarrow q = \frac{2-2p-a}{1+a}$$

   Then, the probabilities are substituted and the equations solved:

$$p = \frac{1-aq}{1+a} = \frac{1 - a\frac{2-2p-a}{1+a}}{1+a} = 1 - \frac{a}{1+a^2}$$
$$q = \frac{2-2p-a}{1+a} = \frac{2 - 2(1 - \frac{a}{1+a^2}) - a}{1+a} = \frac{a - a^3}{1+a+a^2+a^3}$$

   Because of the symmetric pay-off matrix, it is clear that the expressions for agent B are identical to the ones outlined above for agent A, with the values of p and q replaced by the values of v and w respectively. Therefore, $p = v = 1 - \frac{a}{1+a^2}$ and $q = w = \frac{a-a^3}{1+a+a^2+a^3}$. All probabilities are described as a function of a and can be seen in figure 1a. It is clear to see from the figure that, for all possible values of a, a mixed strategy with all three actions is possible.

4. The knowledge of the NE can help the travelers if they are both assumed to be rational. For each of the NE, the expected value can be calculated. If the agents are indeed rational, they both want to maximize their utility and so will choose a strategy that is a NE and has the highest pay-off.



(a) (b)

Figure 1: (a) The probabilities of selecting actions **1**, **2**, or **3** as a function of a in the discrete travelers' dilemma using a mixed Nash equilibrium. (b) The probabilities of selecting those same actions as a function of a using mixed regret minimization.

For the pure strategies of playing **1**, **2**, or **3**, the respective expected utility is commensurate. For the mixed strategy, the expected utility is

$$\begin{aligned}\mathbb{E}u_A(s_A^*, s_B^*) = \mathbb{E}u_B(s_B^*, s_A^*) &= p + q(1+a) + (1-p-q)(1+a) = p + (1-p)(1+a) \\ &= 1 + a - pa = 1 + a - a(1 - \frac{a}{1+a^2}) = 1 + \frac{a^2}{1+a^2}\end{aligned}$$

Thus, the expected utility using the mixed strategies is bounded by $1 + \frac{0}{1+0} = 1 < \mathbb{E}u_A(s_A^*, s_B^*) < 1 + \frac{0.5^2}{1+0.5^2} = 1.2$ and is therefore worse than two of the available pure strategies. If the agents want to maximize their utility, given their knowledge of all the NE, they will both choose the pure strategy of always playing action **3**. It should be noted that this only works because of the symmetry of the problem and because a is set so low that defecting does not make sense.

5. The regret table can be seen below.

A \ **B**	**1**	**2**	**3**
1	0, 0	1-a, a	2-a, a
2	a, 1-a	0, 0	1-a, a
3	a, 2-a	a, 1-a	0, 0

6. Taking the maximum regret for agent A, the strategies **1**, **2**, and **3** yield $2-a$, $1-a$, and a respectively. Minimizing over these values the regret-minimizing strategy is acquired as $s_A = 3$. The exact same process gives the same result for agent B.

   There is also a mixed strategy for regret minimization. Similarly as in question 1, the regret value expectations are calculated per action and solved for equivalency. This yields $p = v = \frac{2a}{3+6a-4a^2}$ and $q = w = \frac{6a+8a^2-8^3}{3+12a+8a^2-8a^3}$. Figure 1b shows graphically what the probabilities of the different actions look like as a function of a.

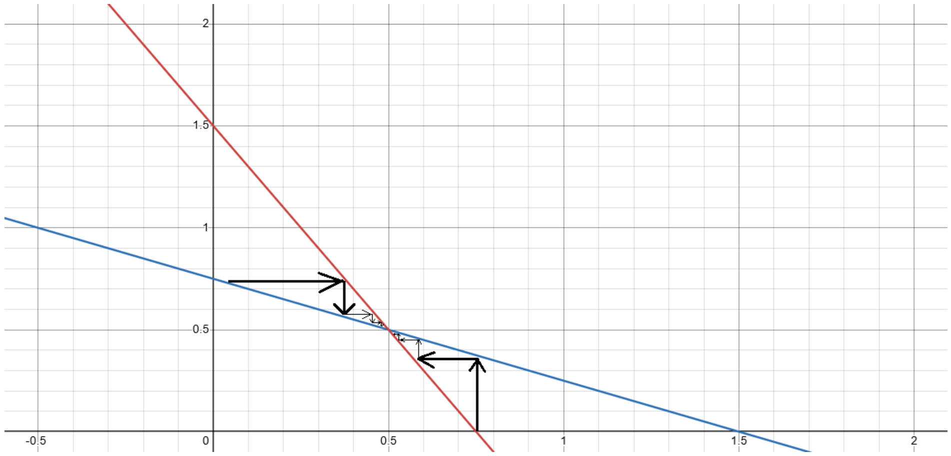


Figure 2: Model of the best responses in Cournot's Duopoly. The model parameters are $\alpha = 2$, $\beta = 1$, $c_1 = c_2 = 0.5$. The best responses made are represented with arrows. No matter where the production is started for either company, if the best responses are followed, the NE point is reached, which in this case is $(0.5, 0.5)$.

1.3 Cournot's Duopoly (continuous version)

1. The best response can be found by setting the derivative of the utility function to 0. First, the unit utility is $p_i(q_i, q_{-i}) - c_i$. Then, the total utility is $u_i(q_i, q_{-i}) = q_i(\alpha - \beta(q_i + q_{-i}) - c_i)$. The following derives the maximum point of q_i:

$$\frac{\partial u_i(q_i, q_{-i})}{\partial q_i} = 1 \cdot (\alpha - \beta(q_i, q_{-i}) - c_i) - \beta q1 = \alpha - 2\beta q_i - \beta q_{-i} - c_1 = 0$$

$$\Rightarrow 2\beta q_i = \alpha - \beta q_{-i} - c_i \Rightarrow q_i = \frac{\alpha - \beta q_{-i} - c_i}{2\beta}$$

   Therefore, the best response for company 1 is $q_1 = \frac{\alpha - \beta q_2 - c_1}{2\beta}$ and for company 2 is $q_2 = \frac{\alpha - \beta q_1 - c_2}{2\beta}$.

2. Figure 2 shows graphically how following the best responses leads to an equilibrium point. The model parameters α, β, c_1, and c_2 only change the intersects, not the underlying properties of the model and thus were chosen arbitrarily as $\alpha = 2$, $\beta = 1$, $c_1 = c_2 = 0.5$. If the first company chooses $q_1 = 0.05$, the best response for company 2 is $q_2 = \frac{\alpha - \beta q_1 - c_2}{2\beta} = \frac{2 - 0.05 - 0.5}{2} = 0.725$. Then company 1 can reply in turn with $q_1 = \frac{\alpha - \beta q_2 - c_1}{2\beta} = \frac{2 - 0.725 - 0.5}{2} \approx 0.388$, and so on. This follows the trajectory from the left, outlined in figure 2. The final point, $(0.5, 0.5)$ in this case, is a Nash equilibrium point as neither company has an incentive to deviate. The best response for both companies here is to not change production $q_1 = q_2 = \frac{2 - 0.5 - 0.5}{2} = 0.5$.

1.4 Ice Cream Time!

Questions:

1. Because of the linearity of the problem, it does not actually matter where Charlize sets her stall, as long as it is between the other two vendors. This can be formalized by saying that the utility of Charlize is constituted from the clientele between the midpoints of her and Alice, and her and Bob, $u_C(c, a, b) = \frac{c+b}{2} - \frac{c+a}{2} = \frac{b-a}{2} = 0.35$, where c is the location of Charlize's stall. It can be seen that the final expression does not depend on c. In this case, the utility of setting up between

the other merchants is better than the utility of setting up on the edges where she could at most get a utility of 0.1 or 0.2 if she placed herself before Alice or after Bob respectively.

2. As explained above, if Charlize sets up between the other vendors, her position does not matter and utility is defined as $u_C(c,a,b) = \frac{b-a}{2}$. So c can be any position x as long as $a < x < b$. If she sets up beyond Bob, her utility is $u_C(c,b) = 1 - \frac{c+b}{2}$. Because Charlize is assumed rational, to maximize her utility in such a case she would set up as close as possible to Bob, $c = b + dc$ with dc as an infinitesimal distance. She should pick this action when $1 - b - \frac{dc}{2} > \frac{b-a}{2} \Rightarrow b < \frac{2+a-dc}{3} \approx 0.7$. Hence, Charlize's best response is:

$$c = \begin{cases} a < x < b, & \text{if } b \geq 0.7 \\ b + dc, & \text{otherwise} \end{cases}$$

3. As worked out in the previous part, if $b < 0.7$, Charlize will establish herself just after Bob, making his utility $u_B(b,a) = b - \frac{b+a}{2} = \frac{b-a}{2}$. In this case, the larger the b the more utility Bob gets, so he should stay as close to 0.7 as possible.

   If $b \geq 0.7$, Bob's utility is $u_B(b,c) = 1 - \frac{b+c}{2} = 1 - b + \frac{b-c}{2}$. In this case, it is clear to see that the smaller the b, the better Bob's utility. Therefore, he should choose $b = 0.7$. Then, from the first term, $1 - b$, Bob is already guaranteed the same utility that he could maximally get if $b < 0.7$. Since all locations between a and b are equivalent for Charlize, Bob cannot guarantee anything from the second term $\frac{b-c}{2}$. But $b = 0.7$ is at least as good any other location, and it is better in expectation if Charlize picks her spot randomly between a and b.

4. Working backwards can yield Alice her optimal stall placement. Begin by assuming that Alice and Bob have already placed their carts optimally at a and b respectively, and that $a < b$. Charlize then must place her cart in one of three places: before a, between a and b, or after b. She will choose one so as to maximize her utility, which will be $u_C(c,a,b) = a$, $u_C(c,a,b) = \frac{b-a}{2}$, and $u_C(c,a,b) = 1 - b$ for the aforementioned scenarios respectively.

   Next, Bob can use this information to optimally place his stall. If he divides the range between a and 1 with a ratio of $2:1$, both sides of his stall will yield equal utility for Charlize. Bob can then nudge his cart infinitesimally away from a so as to guarantee that Charlize sets her cart between a and b, which is preferred for the reasons outlined in the previous subquestion.

   Finally, Alice can now reason that Charlize will place her stall before Bob's, and that she needs to similarly split the space between 0 and b with a ratio of $1:2$. Thus, the final length of the beach will be divided into three parts with lengths $a : b - a : b$ with a ratio of $1:2:1$. Thus, the optimal location for Alice to put her cart is $a = \frac{1}{4}$, or more precisely, it is infinitesimally less than 0.25. By symmetry, she could also place her stall at infinitesimally more than 0.75. The expected value, if Charlize on average puts her cart at the centermost location between a and b, is $\mathbb{E}u_A(a,b,c) = \mathbb{E}\frac{a+c}{2} = \frac{0.25+\mathbb{E}c}{2} = \frac{0.75}{2} = 0.375$.