Homework-1, Multi Agent Systems

Ankur Satya, Azhar Shaikh, Björn van der Haas

November 16^{th} 2022

1 Game Theory: Optimality Concepts and Nash Equilibrium

1.1 Iterated Best response dynamics for a simple matrix game

1. There is no dominated strategy for the given payoff matrix since for a strategy A to be completed by another strategy B, the payoff for either of the players for strategy A should be greater than the payoff of strategy B irrespective of the decision of the other player.

2. For Player 1:
   Best Response(P2=L): M; BR(P2=C): M; BR(P2=R): U

   For Player 2:
   Best Response(P1=U): L; BR(P1=M): C; BR(P2=R): C

   To simulate the best response dynamics, we arbitrarily allow any player to start. Let's consider that player 1 starts first then the following sequences show how the game proceeds for different initial decisions of Player 1:

   P1(U) → P2(L) → P1(M) → P2(C) → P1(M)

   P1(M) → P2(C) → P1(M) → P2(C)

   P1(D) → P2(C) → P1(M) → P2(C)

   For Player 2:
   P2(L) → P1(M) → P2(C) → P1(M)

   P2(C) → P1(M) → P2(C)

P2(R) $\rightarrow$ P1(U) $\rightarrow$ P2(L) $\rightarrow$ P1(M) $\rightarrow$ P2(C) $\rightarrow$ P1(M)

So we see that no matter who starts first and whatever their first move is, the best response dynamics always converges to player 1 choosing M and player 2 choosing C as their moves.

3. If the best response dynamics is applied to a zero sum like matching pennies, the outcome would be completely different qualitatively. It would be a cyclical outcome where the two players would never converge to the mutual best option unlike the given problem.

1.2 Travelers' dilemma: Discrete version

1. The payoff matrix is:

		Player 2		
		1	2	3
	1	$1, 1$	$1+a, 1-a$	$1+a, 1-a$
Player 1	2	$1-a, 1+a$	$2, 2$	$2+a, 2-a$
	3	$1-a, 1+a$	$2-a, 2+a$	$3, 3$

2. We will use the best response dynamics to find Pure Nash equilibrium/equilibria.

For Player 1
BR(P2=1): 1; BR(P2=2): 2; BR(P3=3): 3

For Player 2
BR(P1=1): 1; BR(P1=2): 2; BR(P1=3): 3

So we see that there are 3 Pure Nash equilibria(as they have 3 best mutual response) and they are: Strategies (1,1), (2,2) and (3,3).

3. We assume that:
$P_1(1) = p; P_1(2) = q; P_1(3) = 1 - p - q$
$P_2(1) = l; P_1(2) = m; P_1(3) = 1 - l - m$
where $P_i(t)$ is the probability of taking action t by player i.

To find mixed strategy Nash equilibrium, we will equate the expectation of payoffs for different actions for a player i.e

$$E[P1 = 1] = E[P1 = 2] = E[P1 = 3] \tag{1}$$

$$E[P2 = 1] = E[P2 = 2] = E[P2 = 3] \tag{2}$$

For player 1:

$$E[P1 = 1] = l * 1 + m * (1 + a) + (1 - l - m) * (1 + a)$$
$$E[P1 = 2] = l * (1 - a) + m * 2 + (1 - l - m) * (2 + a)$$
$$E[P1 = 3] = l * (1 - a) + m * (2 - a) + (1 - l - m) * 3$$

Solving for l and m using the above equations and equation 1, we get:

$$l = \frac{a^2 - a + 1}{a^2 + 1}$$
$$m = \frac{-(a - 1) * a}{a^2 + 1}$$

For player 2:

$$E[P2 = 1] = p * 1 + q * (1 + a) + (1 - p - q) * (1 + a)$$
$$E[P2 = 2] = p * (1 - a) + q * 2 + (1 - p - q) * (2 + a)$$
$$E[P2 = 3] = p * (1 - a) + q * (2 - a) + (1 - p - q) * 3$$

Solving for p and q using the above equations and equation 2, we get:

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

4. $p > q > (1 - p - q)\ \forall a : 0 < a < 1/2$
   $l > m > (1 - l - m)\ \forall a : 0 < a < 1/2$

   So, the travelers could make decisions keeping this probability distribution in mind but then again it doesn't exactly tell them which decision is the best rather it just gives them a comparative advantage.

5. The regret matrix is:

		Player 2		
		1	2	3
Player 1	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. Regret minimization strategies are given by:

$$s_i^{rm} = argmin_{s_i} R_i^{max}(s_i)$$
$$R_i^{max}(s_i) = max_{s_j} R_i(s_i, s_j)$$

Using the regret matrix we find:

$R_1^{max}(s_1) = 2 - a$; $R_1^{max}(s_2) = 1 - a$; $R_1^{max}(s_3) = a$

$R_2^{max}(s_1) = 2 - a$; $R_2^{max}(s_2) = 1 - a$; $R_2^{max}(s_3) = a$

So,

$$s_1^{rm} = 3$$
$$s_2^{rm} = 3$$

So the regret minimization strategy is (s_1, s_2): (3, 3)

1.3 Cournot's Duopoly (continuous version)

1. Unit price of the bottle in the market is given by $p(q_1, q_2) = \alpha - \beta(q_1 + q_2) \quad (\alpha, \beta > 0)$.
   Now let's consider company 1.
   total cost $= c_1 q_1$
   profit $= pq_1 - c_1 q_1 = (\alpha - \beta(q_1 + q_2) - c_1)q_1$
   The best response of each company is to maximise their individual profits. This entails the optimal q_1^* can be found by taking the partial derivative of the profit w.r.t to q_1 at some given value q_2 and setting to zero.

$$\frac{\partial(profit)}{\partial q_1} = \frac{\partial(\alpha q_1 - \beta q_1^2 + \beta q_2 - c_1 q_1)}{\partial q_1} = 0$$

$$= \alpha - 2\beta q_1 + \beta q_2 - c_1 = 0 \equiv q_1^* = \frac{\alpha + \beta q_2 - c_1}{2\beta}$$

   Therefore the best response of company 1 is to produce q_1^* quantity. Similarly the best response of company 2 is given by q_2^*

$$q_2^* = \frac{\alpha + \beta q_1 - c_2}{2\beta}$$

2. The companies will react to each other until they reach the nash equilibrium(NE). As illustrated in the diagram below.

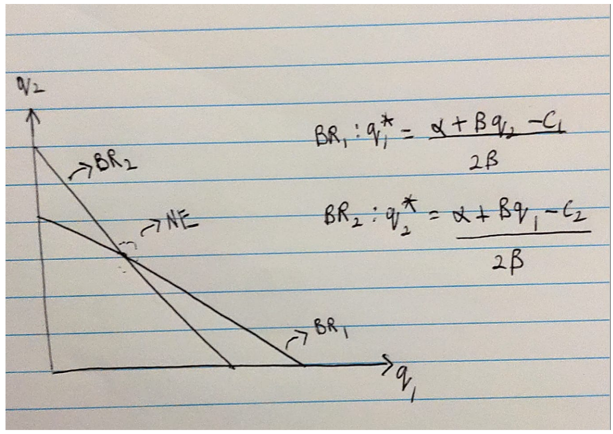

Figure 1: Nash Equilibrium

1.4 Ice cream time!

1. This situation is a continuous (infinite) action space. We can calculate the utility with only Alice and Bob present as followed: Ua(a, b) = a + ((b-a)/2) = 0.1 + ((0.8-0.1)/2) = 0.45 and Ub(a, b) = 1 – ((a+b)/2) = 1 – ((0.1+0.8)/2) = 0.55.

   Given this, Charlize could position herself in the middle of both Alice and Bob, (c = 0.45), giving her Uc (a,b,c) = (Ua/2) -a + (Ub/2) – (1 – b) = 0.2. Resulting in Ua = 0.45 – (Uc / 2) = 0.35 and Ub = 0.55 – (Uc / 2) = 0.45. If Charlize positions herself to the right of Bob (0.81), she would obviously get less than 0.2.

   If Charlize positions herself to the right of Alice or left of Bob (c = 0.11 or c = 0.79 or an even smaller difference respectively) then Uc = Ua – a = 0.35 (or Uc = Ub – (1 – b) = 0.35). This is Charlize's Best Response. However, because Alice and Bob can also move they would eventually each move to the right and left of Charlize leaving her with almost no utility. Thus, assuming Alice and Bob can still move she would be better off moving to the right of Bob which eventually results in Bob having little to no utility (as Bob would move closer to Alice, allowing Charlize to move closer to Bob).

2. As explained in 1), her best move would be to the right of Bob.

3. Bob should move as close to the right as possible (nearing 1). This forces Charlize to position herself somewhere in the middle, allowing Bob (and Alice) to eventually move closer to Charlize and gain even more utility.

4. If she immediately puts herself at location 0.1, she gives the next person (Bob or Charlize) the chance to position themselves right next to Alice which results in Alice having little to no utility. Instead, Alice should position herself slightly to one side of the middle (0.45 or 0.55) so that the next person is incentivized to move to the opposite side of the middle and temporarily have the most utility. The third person that arrives is then incentivized to put themselves next to the second person to optimize their utility leaving the second person with almost no utility.