UNIVERSITY OF TENNESSEE, KNOXVILLE
Department of Mathematics
MATH123 Finite Mathematics Section 8
Written by Calvin WONG

April 28, 2023: Optimal Mixed Strategy

Reminder

1. Learning Activities' scores are uploaded to WebAssign regularly. Please regularly check if we have inputted your score correctly.
2. WebAssign Lesson 18 due on Sunday, Apr 30.

Recaps

1. For every matrix game, it can only be either strictly determined or non-strictly determined. For a strictly determined game, pure strategies (maximin for row player and minimax for column player) can be used to solve the game. In such game, both players are satisfied by the outcome.
2. If a game is not strictly determined, then no pure strategies exist, and we need to use mixed strategies, which is a class of strategies that utilize probabilities.
3. If a game with the 2×2 payoff matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is not strictly determined, then we know $a+d-b-c \neq 0$. Furthermore, if the optimal mixed strategies of row player and column player are $P = \begin{bmatrix} p_1 & 1-p_1 \end{bmatrix}$ and $Q = \begin{bmatrix} q_1 \\ 1-q_1 \end{bmatrix}$, we have

$$p_1 = \frac{d-c}{a+d-b-c}, \tag{1}$$

$$q_1 = \frac{d-b}{a+d-b-c}. \tag{2}$$

   Also, the game value (the expected value of the outcome for the row player) is given by

$$E = PAQ = \frac{ad-bc}{a+d-b-c} \tag{3}$$

4. The formula above is listed in the formula sheet. They are formula I, J and K respectively.
5. Before using formula I, J, K, *you must check the game is not strictly determined first.*

Warm up Question

Raymond and Calvin are playing an altered Rock-Paper-Scissors. Raymond is only allowed to play rock and paper, and Calvin is only allowed to play rock and scissors. Whoever loses the game needs to pay their opponent $1.

The game matrix is given as follows.

$$\begin{array}{r c l}
 & \begin{array}{cc} \text{Rock} & \text{Scissors} \end{array} & \text{min} \\
\begin{array}{r} \text{Rock} \\ \text{Paper} \end{array} & \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} & \begin{array}{l} 0 \\ -1 \end{array} \\
\text{max} & \begin{array}{cc} 1 & \quad 1 \end{array} &
\end{array}$$

As we can check, the maximin value is 0 and the minimax value is 1. So, the game is not strictly determined. No pure strategies exist for both players. Using the given formula, we have $a + d - b - c = 0 - 1 - 1 - 1 = -3$, and

$$p_1 = \frac{d-c}{a+d-b-c} = \frac{(-1)-1}{-3} = \frac{2}{3},$$

$$q_1 = \frac{d-b}{a+d-b-c} = \frac{(-1)-1}{-3} = \frac{2}{3},$$

$$E = PAQ = \frac{ad-bc}{a+d-b-c} = \frac{(0)(-1)-(1)(1)}{-3} = \frac{1}{3}.$$

In conclusion,

1. In every 3 consecutive games, Raymond should play Rock twice and Paper once.
2. In every 3 consecutive games, Calvin should play Rock twice and Scissors once.
3. While both players play their respective optimal strategies, the expected payoff for Raymond is $\$\frac{1}{3}$. Thus, the game is favourable to Raymond.

Question: How does the formula works?

Let's put ourselves into Raymond's shoes. Assume Raymond is rational, and he has done all the analysis, and acknowledged that the game is not strictly determined. Now, he will determine a mixed strategy for himself. The only problem is, he don't know what value of p he should pick.

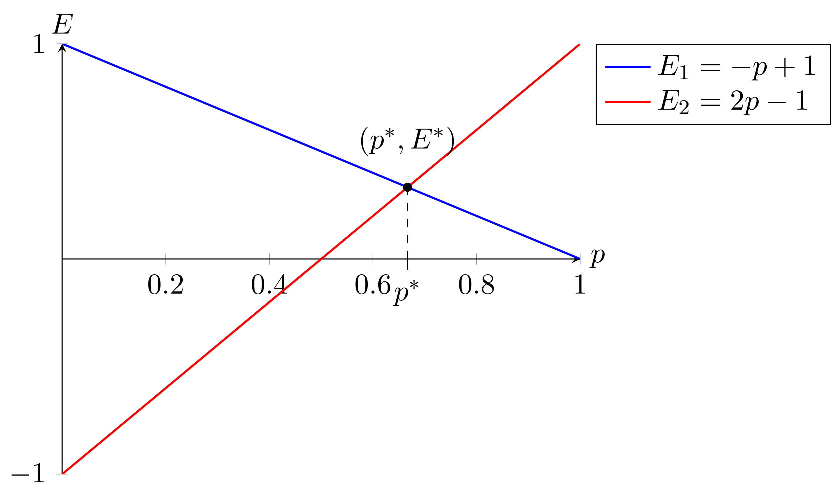
$$\begin{array}{ll} & \begin{array}{cc} \text{Rock} & \text{Scissors} \end{array} \\ \begin{array}{lr} \text{Rock:} & p \\ \text{Paper:} & 1-p \end{array} & \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \end{array}$$

Now, Calvin has two options, playing Rock or playing Scissors. Let's try to find the expected outcome for Raymond of each case.

1. If Calvin plays Rock, the expected payoff of Raymond is $E_1 = (p)(0) + (1-p)(1) = -p + 1$.
2. If Calvin plays Scissors, the expected payoff of Raymond is $E_2 = (p)(1) + (1-p)(-1) = 2p - 1$.

If you are not familiar to the concept of expected value, an easy way to understand the above calculations is $PA = [-p+1 \;\; 2p-1]$.

We can plot the two equations on a $p - E$ plane.



Note that the two lines intersect at a point (p^*, E^*). We can consider the following three situations.

<u>Case 1: Raymond choose p so that $p < p^*$</u>

In this case, Calvin will always play Scissors, since the red line is lower than the blue line. In this way, Calvin can minimize the payoff to Raymond. In particular, Calvin has a choice to switch his strategy to gain more, and so Raymond's strategy this time is not a optimal strategy.

<u>Case 2: Raymond choose p so that $p > p^*$</u>

Similar to the last case, and Calvin will always play Rock.

Case 3: Raymond choose p so that $p = p^*$

As observed in the graph, the payoff for Raymond is the same when Calvin plays either Rock and Scissors. In this case, Calvin cannot gain additional benefit from switching options. Hence, this is the optimal strategy for Raymond. Also, from solving the system of linear equations, we have $(p^*, E^*) = \left(\frac{2}{3}, \frac{1}{3}\right)$, which is the same as using the formula.

We can repeat the above thinking process for Calvin to find his optimal strategy too, and it is left as an exercise for the readers.

Appendix - Proof of the Formula

Let $F = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a game matrix with no saddle points and $a + d - b - c \neq 0$.
Show that

1. maximin strategy for the row player is $\vec{x} = \left(\dfrac{d-c}{a+d-b-c}, \dfrac{a-b}{a+d-b-c} \right)$,

2. minimax strategy for the column player is $\vec{y} = \left(\dfrac{d-b}{a+d-b-c}, \dfrac{a-c}{a+d-b-c} \right)$,

3. game value $E = \dfrac{ad-bc}{a+d-b-c}$.

Answer.

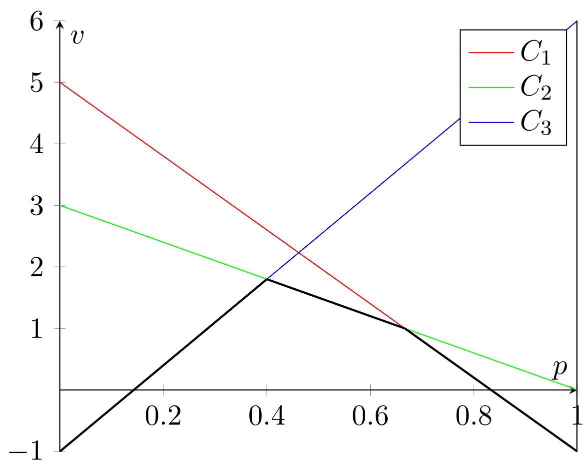
If Column player choose C_1, payoff of Row player $= (a)(p) + (c)(1-p) = (a-c)p + c$;
if Column player choose C_2, payoff of Row player $= (b)(p) + (d)(1-p) = (b-d)p + d$.
So, in order not to give any advantages to the Column player,
we must have $(a-c)p + c = (b-d)p + d$.
So, $p = \dfrac{d-c}{a-b-c+d}$ and the optimal mixed strategy for Row player
will be $\left(\dfrac{d-c}{a-b-c+d}, 1 - \dfrac{d-c}{a-b-c+d} \right) = \left(\dfrac{d-c}{a-b-c+d}, \dfrac{a-b}{a-b-c+d} \right)$.
Then, the game value = payoff to the Row player
$=(a-c)\left(\dfrac{d-c}{a-b-c+d} \right) + c = \dfrac{ad - ac - cd + c^2 + ac - bc - c^2 + cd}{a-b-c+d} = \dfrac{ad-bc}{a-b-c+d}$.
If Row player choose R_1, payoff of Row player $= (a)(q) + (b)(1-q) = (a-b)q + b$;
if Row player choose R_2, payoff of Row player $= (c)(q) + (d)(1-q) = (c-d)q + d$.
So, in order not to give any advantages to the Row player,
we must have $(a-b)q + b = (c-d)q + d$.
So, $q = \dfrac{d-b}{a-b-c+d}$ and the optimal mixed strategy for Column player
will be $\left(\dfrac{d-b}{a-b-c+d}, \dfrac{a-c}{a-b-c+d} \right)$.

Appendix - An Example with a 2×3 Payoff Matrix

For the following games, assume $\vec{x} = (p, 1-p)$ and $\vec{y}$ as the mixed strategies of the row and column player respectively.
It is given the game given have no saddle point. Find $\vec{x}$, $\vec{y}$ and the game value.

$$\begin{bmatrix} -1 & 0 & 6 \\ 5 & 3 & -1 \end{bmatrix}$$

Answer.
If Column player choose C_1, payoff of Row player $= (-1)(p) + (5)(1-p) = -6p + 5$;
if Column player choose C_2, payoff of Row player $= (0)(p) + (3)(1-p) = -3p + 3$;
if Column player choose C_3, payoff of Row player $= (6)(p) + (-1)(1-p) = 7p - 1$;
We then draw the three payoff line and the lower envelop (the black curve) on the same plane:



So, we can see that the intersection of C_2 and C_3 will be the optimal.
By solving $\begin{cases} v = -3p + 3 \\ v = 7p - 1 \end{cases}$, we have $(p, v) = (0.4, 1.8)$.
So, the optimal mixed strategy of row player is $(0.4, 0.6)$ and the game value is 1.8 .
Now, note that C_2 and C_3 are the two relevant options. So, we let $\vec{y} = (0, r, 1-r)$.
By solving $\begin{cases} (0)(r) + (6)(1-r) = 1.8 \\ (3)(r) + (-1)(1-r) = 1.8 \end{cases}$, we have $r = 0.7$.
So, the optimal mixed strategy of column player is $(0, 0.7, 0.3)$.

Appendix - Wrongly applying formula I,J,K to strictly determined game

Considering the following payoff matrix.

$$\begin{array}{ccc} & & \text{min} \\ & \begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix} & \begin{array}{c} 1 \\ 3 \end{array} \\ \text{max} & \begin{array}{cc} 3 & 4 \end{array} & \end{array}$$

Since the maximin value is equal to the minimax value, saddle point exists and the game is strictly determined. The maximin strategy of row player is to play row 2, and the minimax strategy of the column player is to play column 1.

Now, let's try to use the formula I,J,K and see what will happen.

$$p_1 = \frac{d-c}{a+d-b-c} = \frac{4-3}{1+4-3-3} = -1,$$

$$q_1 = \frac{d-b}{a+d-b-c} = \frac{4-3}{1+4-3-3} = -1.$$

$$E = PAQ = \frac{ad-bc}{a+d-b-c} = \frac{(1)(4)-(3)(3)}{1+4-3-3} = 5$$

It is clear that having negative value of p_1 and q_1 is an absurd situation.

END OF NOTE